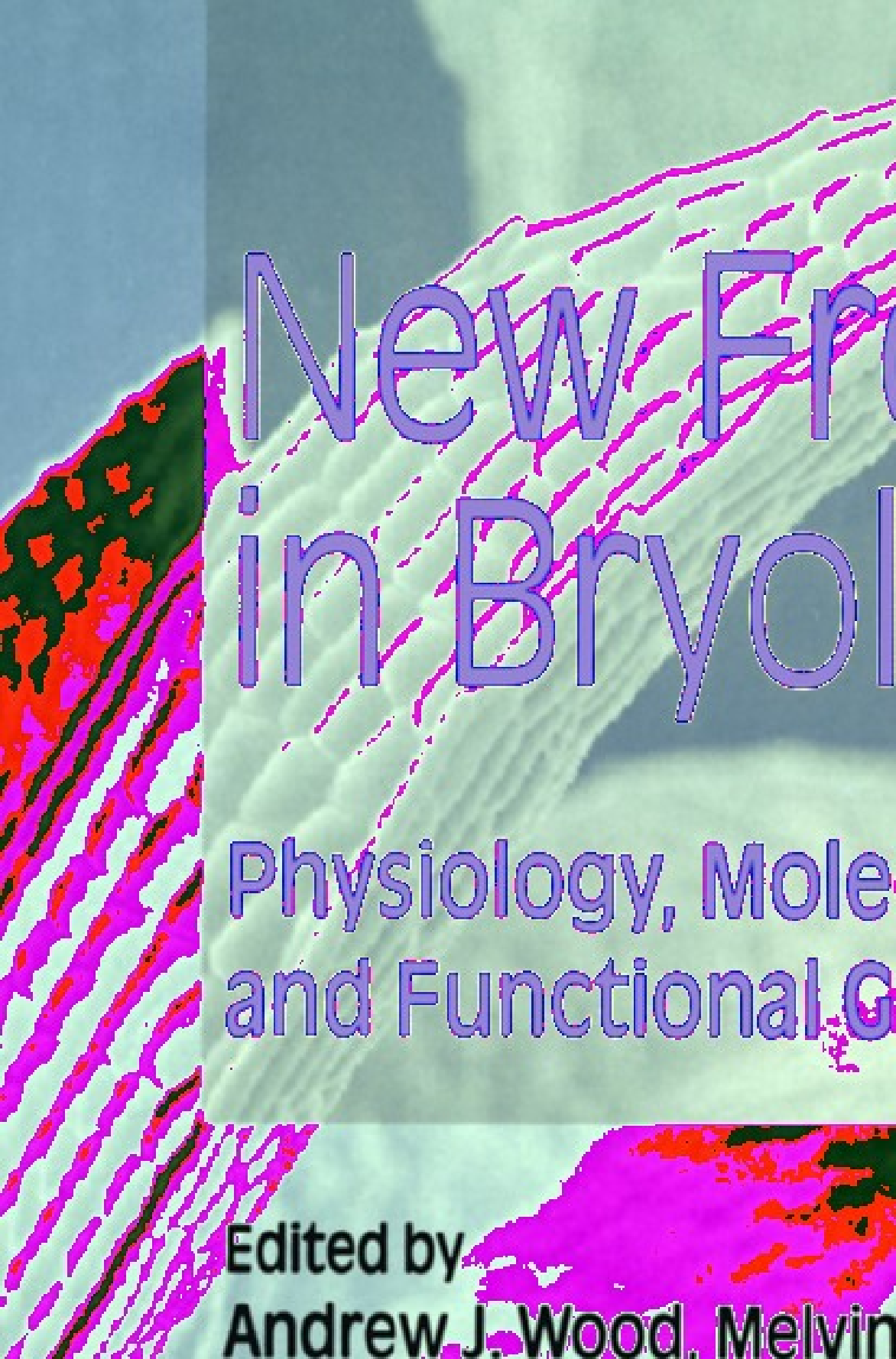
New Fr
in Bryol
Physiology, Mole
and Functional G
Edited by
Andrew J. Wood, Melvin

New Frontiers in Bryology

New Frontiers in Bryology

Physiology, Molecular Biology and Functional Genomics

Edited by

Andrew J. Wood
Southern Illinois University-Carbondale, Carbondale, U.S.A.

Melvin J. Oliver
USDA-ARS, Lubbock, U.S.A.

and

David J. Cove
University of Leeds, Leeds, U.K. and Washington University, St. Louis, U.S.A.

SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-90-481-6569-8 ISBN 978-0-306-48568-8 (eBook)
DOI 10.1007/978-0-306-48568-8

Printed on acid-free paper

Originally published by Kluwer Academic Publishers in 2004
Softcover reprint of the hardcover 1st edition 2004

TABLE OF CONTENTS

PREFACE

The mosses (Bryophatea, Musci) are a diverse and widely distributed group of land plants. Mosses are attractive experimental plants because they exhibit the traditional attributes of good model systems (i.e. ease of growth & maintenance, fast generation time, and amenable genetics) with the added advantage of a haploid gametophyte that allowed developmental mutants to be recovered with relative ease. In addition, mosses with the ability to tolerate extreme environmental conditions offer realistic models for the analysis of environmental stress-tolerance; particularly when compared to tracheophytes such as *Arabidopsis thaliana* in which these important plant phenotypes are either not clearly expressed or entirely lacking. And, in one of the most exciting developments in Plant Biology, efficient homologous recombination occurs in the moss *Physcomitrella patens*. The ability to perform efficient homologous recombination (i.e. gene knock-outs) in *P. patens* is at present unique amongst all plants and represents an extremely powerful technique for the functional analysis of many plant genes.

Over the past 5 years, a world-wide community of moss researchers has evolved. A highly successful "Moss" conference has been held annually (1998—Mumbai, India; 1999—Carbondale, IL, USA; 2000—Villars, Switzerland; 2001—Okazaki, Japan; 2002—Ambleside, UK; 2003—St. Louis, MO, USA) with "Moss 2004" planned to be held in Friéburg Germany. These conferences have been instrumental in the creation & development of strong collaborative ties, and the free exchange of both ideas and materials. Mosses are powerful experimental tools for the elucidation of complex biological processes in plants—from evolution & development to the homologus recombination & the creation of novel molecular tools. This book provides a synopsis of the outstanding *basic* research being conducted using mosses as a model multi-cellular eukaryote.

Finally, we gratefully acknowledge Professor Govindjee (Emeritus, Department of Plant Biology, University of Illinois at Urbana-Champaign) for encouraging the production of a volume devoted to recent advances in bryology.

Andrew J. Wood
Melvin J. Oliver
David J. Cove
October 15th, 2003

COLOUR SECTION

1)

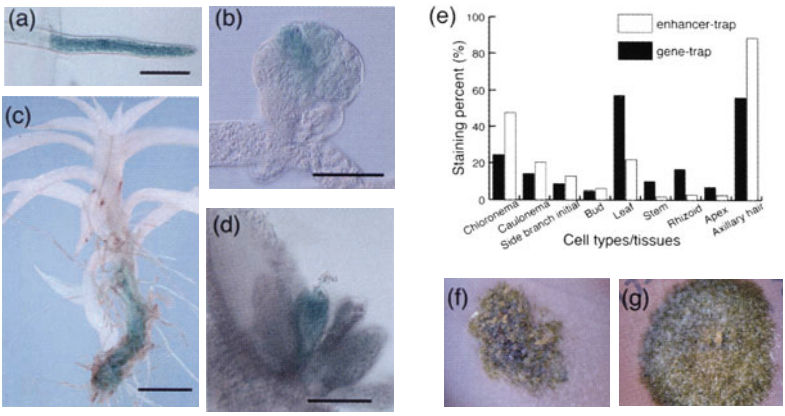

Figure 2. GUS expression patterns of gene-trap and enhancer-trap lines, and a tagged mutant with an altered hormone response. Staining of a rhizoid apical cell of gene-trap line YH87. (b) A bud of a young gametophore of gene-trap line YH229, showing staining of an apical cell and its surrounding cells. (c) A gemetophore of gene-trap line YH209, showing staining of the basal region of the gametophore, where rhizoid filaments emarged. (d) A shoot apex of gene-trap line YH440, showing staining of an antheridium. Bars in (a) and (b) = 50 μ*m, in (c) = 500* μ*m, and in (d) = 100* μ*m. (e)* Patterns of expression of the uidA gene in gene-trap and enhancer –trap lines generated with the homologus integration method. Y-axis values indicate the percentage of lines with expression in the indicated portions relative to the number of GUS-positive gametophytes, which was set at 100%. Apex indicates the number of lines displaying GUS staining in an apical cell and its surrounding cells. (f) A month-old wild type gametophyte grown on medium supplemented with 1 mM BA, showing malformed bud formation. (g) A month-old gametophyte of tagged line 8617-7, which is resistant to cytokinin, grown on the same medium as the wild type. (see Fig. 2. on p. 117)

2)

Figure 1. Batch culture of Ceratodon purpurenus in white light (*70 W/m^{-2}; 16h light/8 h dark; 20^{o}C). The colour changes form deep green to brownish during the culture of 21 days. Cells of this type are full of oil drops like in figure 2.* *(see Fig 1. on p. 136)*

3)

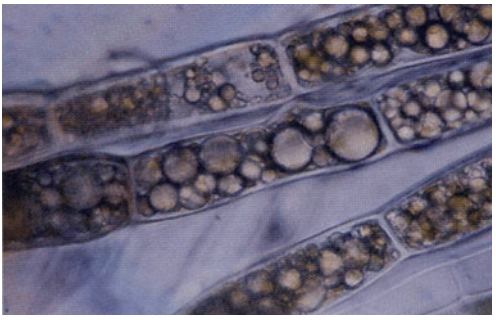

Figure 2. Senescent protonema filaments of Ceratodon purpureus from 21 day-old batch culture. *The cells are full of oil drops.* (see Fig. 2 on p. 137)

MOLECULAR PHYLOGENY OF BRYOPHYTES AND PECULIARITIES OF THEIR CHLOROPLAST AND MITOCHONDRIAL DNAS

VOLKER KNOOP, YIN-LONG QIU, KOICHI YOSHINAGA

Institut für Zelluläre und Molekulare Botanik, Abt. Molekulare Evolution,Universität Bonn, Kirschallee 1, D-53115 Bonn, Germany (VK); Department of Ecology and Evolution, 830 North University Ave., University of Michigan, Ann Arbor, MI 48109-1048(YQ); Shizuoka University, Faculty of Science, Oya 836, 422-8529 Shizuoka, Japan(KY)

Abstract. Molecular sequence data have contributed enormously to our knowledge about the phylogeny of land plants. Bryophytes are of fundamental importance to the full picture of land plant evolution as it is becoming increasingly evident that they represent the extant relatives of the earliest land plants. The three classically distinguished bryophyte classes mosses, liverworts and hornworts each are confirmed as well founded monophyletic groups by the majority of molecular data. However, it has also become clear that the term bryophytes describes a division, which is a paraphyletic group – most likely with only one of its classes linked to the vascular plants and another one sister to all other land plants. Hitherto enigmatic genera like *Takakia* or *Haplomitrium* can now be phylogenetically assigned and more detailed placements of these and other genera will be straightforward with yet more informative data accumulating. The use of concatenated alignments will likewise resolve other questions on the monophyly of orders or families defined by classical systematics and will paint a clear picture on their branching pattern. Extending the conventionally used nuclear gene sequences (mostly ribosomal RNA genes) or chloroplast gene markers, the plant mitochondrial genes appear to be of particular suitability to address questions of deep level plant phylogeny. Two idiosyncratic phenomena of gene expression in plant organellar DNAs make their evolution particularly interesting: RNA editing and Trans-Splicing. RNA editing, the frequently observed site-specific exchange of pyrimidines to reconstitute conserved codon identities is class-specifically shaped in the bryophytes: Moderate exchange of cytidine to uridine is observed in mosses – very similar to angiosperms, where the phenomenon was discovered. Hornworts, however, display a strikingly higher frequency of pyrimidine exchange and, most notably, many replacements in the opposite direction, i.e. from uridine to cytidine. A peculiar case are the liverworts where we observe RNA editing of variable frequency among taxa in the jungermanniid species but so far a mysterious absence of RNA editing in the marchantiid mitochondrial sequences. Two other genomic features make the mitochondrial DNA of plants particularly well suited for deep level phylogenetic analyses: A very low primary sequence drift coinciding with a low degree of homoplasy and the presence of characteristic, positional stable organellar introns, mostly of the group II type. Investigating the intron-rich genes nad5, nad2, nad4 and nad7 one finds that intron presence fully coincides with class membership, with the mosses displaying spermatophyte-type mitochondrial introns. These and other observations so far converge on the liverworts as sister group to all other land plants.

1. INTRODUCTION

1.1. The bryophytes.

Bryophytes are key to understanding the evolution of land plant life, because data from different sources, including cladistic analyses of molecular sequences, converge on the view that bryophyte-like organisms were the first plants on land, emerging more than 450 million years ago (Kenrick & Crane, 1997; Bremer,

A.J. Wood et al. (ed.),
New Frontiers in Bryology: Physiology, Molecular Biology and Functional Genomics, 1–16.

Humphries, Mishler, & Churchill, 1987). Alternative claims that view bryophytes as a degeneration of early vascular plants have only rarely received support from molecular observations (Yoshinaga, Kubota, Ishii, & Wada, 1992) and never from any cladistic analyses of molecular data of adequate taxon sampling. The early diversification of the division bryophytes into its classes mosses, hornworts and liverworts may have been a very early event of land plant (embryophyte) evolution and this, together with long-branching algal outgroups, are most likely the fundamental reasons why a fully resolved phylogeny of earliest land plants is still lacking. However, the emerging molecular consensus stands in contrast to the disappointing lack of bryophyte macrofossil documents (Edwards, Duckett, & Richardson, 1995; Edwards, 2000; Edwards, Wellman, & Axe, 1998). Only the microfossil record gives evidence for early land plants with a bryophyte-level of organization in the mid-Ordovician (Wellman & Gray, 2000; Graham, 1996), more specifically of land plant spores with a liverwort-like morphology (Taylor, 1995).

It is mostly agreed upon that the origin of land plant life has to be searched among extinct relatives of the Charophyte algae (Graham, 1996; Graham, Cook, & Busse, 2000; McCourt, 1995), more specifically among members of the order Charales (Malek, Lättig, Hiesel, Brennicke, & Knoop, 1996; Karol, McCourt, Cimino, & Delwiche, 2001). New molecular data, among which the complete organellar sequences of the chloroplast and mitochondrial DNAs (Turmel, Otis, & Lemieux, 2002), the "plastomes" and the "chondriomes", appear most promising, will help to ultimately resolve this issue in the near future.

Given that mosses have frequently served as model organisms in biology in which many fundamental insights have been gathered, see (Reski, 1998a) for examples, and given that *Physcomitrella* has emerged as a new model plant of molecular biology (Reski, 1998b) it is of utmost importance to understand how bryophytes relate to other forms of land plant life in evolutionary terms.

1.2. Cladistics and molecular characters.

More than twenty years ago Miller had already stated that "we have come a long way since Micheli (1729) established the discreetness of several taxa of bryophytes and Linnaeus incorporated mosses and liverworts into the *Species Plantarum* (1753)" (Miller, 1979). From today's retrospective this statement may have come a little early at a time when molecular sequence data had not yet made their way into phylogeny and systematics. His and other articles in the same book (Schuster, 1979) give examples for discussion on "primitive" and "advanced" characters when consequent cladistic argumentation were not yet rigorously enough applied. New insights during the following two decades have mainly come from consequent implementation of cladistic analyses following the trend-setting contributions of Mishler and colleagues (Mishler & Churchill, 1984; Mishler, 1986).

Today it is highly unlikely that morphological, biochemical or physiological investigations alone could provide significant numbers of new characters allowing to answer remaining phylogenetic questions unambiguously and with statistical reliability. Thus, only molecular sequence data can be expected to

resolve pending phylogenetic questions, given that only the wealth of data to be obtained there will provide a sufficient number of characters for phylogenetic analyses. Clearly, molecular and morphological data can then be used in combined analyses (Mishler et al., 1994). However, an important aspect is that any type of character, molecular or morphological, can only be used when strict cladistic rules are applied and this has been well established for the bryological field and beyond with the early papers by Mishler and colleagues.

While generally richer in informative characters, molecular sequences are generally prone to the same pitfalls as any other character sampling, most notably to the effects of homoplasy (multiple nucleotide or amino acid exchanges at the same site re-establishing an original character state). Clearly the rate of sequence divergence of a given nucleotide or protein sequence must correlate with the evolutionary time-frame under investigation and consequently the choice of a genetic locus for a given phylogenetic question cannot be arbitrary. A fast evolving sequence will only suffice to answer questions of short-term phylogeny, for example on family level of angiosperms. Only slow evolving loci will have been able to conserve phylogenetic information on phylogenetic old diversification such as the dichotomies among the earliest land plants. As a recently emerging example, the liverworts provide an interesting example for divergent mutation rates in the three genomes of a plant cell: The mitochondrial nad5 locus can provide resolution of the jungermanniid liverworts but no resolution is obtained among the marchantiid liverworts (Beckert, Steinhauser, Muhle, & Knoop, 1999) while the nuclear 18S rRNA gene (Capesius & Bopp, 1997) presents quite the opposite picture.

Given the progress in lab technologies the ongoing phylogenetic research has seen two trends in recent years: extending data matrices vertically, i.e. investigating increasing number of taxa and/or extending data matrices horizontally, i.e. identifying and exploring further informative loci. The latter approach is fruitful in any case. The discussion on matters of taxon sampling, however, which appears to be less relevant in some (Rosenberg & Kumar, 2001) but more relevant in other cases (Rydin & Källersjö, 2002), is continuing. Certainly, in any case the molecular sequence under investigation must necessarily be of sufficient length to convey a sufficient number of characters that withstand statistical tests, the most generally applied of which is the bootstrap re-sampling procedure. A striking example of phylogenetic information content in a very short gene stretch is the early report of a land plant phylogeny derived from the 5S rRNA (van de Peer, De Baere, Cauwenberghis, & De Wachter, 1990). The general outline of land plant phylogeny deduced from this short gene sequence of only 120 nt. was largely correct except for the suggested monophyly of bryophytes, as we will discuss below.

2. THE MOLECULAR CHARACTERS IN USE

The appropriateness of a locus for a given phylogenetic group in a given taxonomic range can never be reliably predicted and the example of diverging sequence drift among the liverwort subclasses stated above may be the most striking example among bryophytes. Nevertheless, a nuclear gene rich in intron sequences is likely the

better *a priori* choice for a within-family or within-genus phylogeny while more conserved structural RNAs or protein encoding regions in the organellar genomes are likely the better choice for higher taxonomic levels up to classes and divisions. The advent of PCR technology has initially led botanical laboratories to analyse large taxon sets on commonly analysed loci, most notably those introduced early into molecular plant phylogenetics such as the chloroplast rbcL gene (Manhart, 1994; Lewis, Mishler, & Vilgalys, 1997) or the nuclear 18S rRNA (Kranz et al., 1995). Studies on 18S rRNA have led to proposals for a revision in liverwort systematics (Capesius, 1995; Bopp & Capesius, 1995; Capesius et al., 1997) that remained unsupported by other studies, also those based on the 18S rRNA gene as well (Hedderson, Chapman, & Rootes, 1996). We consider it likely that this is a rare example of phylogenetic deductions being strongly dependent on several factors simultaneously: the strategies to obtain indel-rich sequence alignments, the adequate inclusion of key taxa and a potential inclusion of paralogous but not orthologous gene copies for alignment construction (Qiu et al., *unpublished observations*).

Studies, in which phylogenetic analysis among bryophytes have been extended to investigate potential novel phylogenetically informative loci include the chloroplast rps4 gene encoding ribosomal protein S4, first in the moss order Polytrichales (Hyvönen, Hedderson, Smith Merrill, Gibbings, & Koskinen, 1998), and later with a wider sampling of mosses (Goffinet, Cox, Shaw, & Hedderson, 2001). The atpB-rbcL spacer (Chiang & Schaal, 2000) was reported in a phylogenetic study with an as yet somewhat limited taxon sampling, similarly the chloroplast ITS (internal transcribed spacer) of the ribosomal RNA gene cluster (Samigullin et al., 1998). The psaB, psbD, psaA and rpoC2 genes were used in a study together with rbcL, to address questions of bryophyte class relationships to other land plants (Nishiyama & Kato, 1999).

More frequently used are other chloroplast loci, notably the intron-containing trnL and the subsequent trnF tRNA encoding genes (Buck, Goffinet, & Shaw, 2000). Several investigations have combined phylogenetic information from different genomes in the plant cell, e.g. a combination of chloroplast and nuclear 18S rRNA loci (Cox & Hedderson, 1999). Frequently the fusion of chloroplast trnL-trnF, rps4 and rbcL loci has been used to address phylogenetic questions among moss clades such as the haplolepidous mosses ("Dicranidae") (La Farge et al., 2000), the hypnobryalean pleurocarpous mosses (De Luna et al., 2000) or the diplolepidous-alternate mosses ("Bryidae"), the latter in fusion with the nuclear 18S rRNA (Cox, Goffinet, Newton, Shaw, & Hedderson, 2000). Occasionally, interesting observations are made when single molecular datasets seem to give contradicting phylogenies. Reticulate evolution, i.e. inter-lineage hybridisation, is postulated to explain divergences in section assignment of the peatmosses (genus *Sphagnum*) when nuclear and chloroplast markers are compared (Shaw & Goffinet, 2000b).

A compilation of 90 molecular systematic studies of bryophyte phylogeny up to the year 2000 was recently published (Goffinet & Hax, 2001). Many contributions were dedicated to specific bryophyte subgroups on different taxonomic levels: families, e.g. the Wardiaceae (Hedderson, Cox, & Gibbings, 1999); orders, e.g. the Orthotrichales (Goffinet, Bayer, & Vitt, 1998); subclasses, e.g. the

Marchantiidae (Wheeler, 2000; Boisselier-Dubayle, Lambourdiere, & Bischler, 2002). Several recent studies have introduced novel molecular marker gene, mostly on lower taxonomic ranges. A new nuclear gene, adenosine kinase, has been explored for the phylogeny of the Amblystegiaceae (Vanderpoorten, Hedenas, Cox, & Shaw, 2002). The ITS1-5.8S-ITS2 rRNA spacer region was employed to address phylogeny in the species rich liverwort genus *Plagiochila* (Heinrichs, Proschold, Renker, Groth, & Rycroft, 2002). The glyceraldehyde-3-phosphate-dehydrogenase has been explored for investigations within the genus *Mitthyridium* (Calymperaceae), where expectedly phylogenetic resolution is derived from variations in the intron sequence of this nuclear protein coding gene (Wall, 2002).

3. MITOCHONDRIAL DNA IN PLANTS

The mitochondrial DNA has appeared as a relative latecomer in land plant phylogeny and this may have been due in part to misconceptions originating in views of the compact and fast evolving animal mitochondrial DNA. Plant mitochondrial DNA, however, is different from its animal counterpart in nearly every aspect (Knoop & Brennicke, 2002). In the plant cell, the mitochondrial genome (the "chondriome") carries more genes than in metazoa, carries introns in many of these and also it carries foreign DNA insertions originating from the chloroplast and the nuclear genome. But – and this is crucial for "deep level" phylogenetics – plant mitochondrial DNA shows a significantly slowed primary sequence drift. This characteristic allows identification of sequences which may contain more information of old evolutionary histories. As an additional benefit, mitochondrial DNA is essentially free of problems caused by multiple genes or gene families, commonly present in nuclear DNA, in other words free of the caveat of comparing paralogous, but not orthologous genes.

Compared to chloroplast DNA (the "plastome"), sequence drift is yet slower in the plant chondriome and an additional element of phylogenetic information rests in the presence of mitochondrial introns with a certain degree of variability in presence among land plant clades as discussed in the next section and contrasting the highly conserved ones in the plastome. Indeed, the presence of three mitochondrial introns which are absent in liverworts but present in all other embryophyte clades supports the phylogenetic view placing liverworts as the sister clade to other land plants (Qiu, Cho, Cox, & Palmer, 1998).

Initial investigations on plant mitochondrial DNA started with a short coding region of the cox3 gene (Hiesel, von Haeseler, & Brennicke, 1994) and were later extended to a second region of the same gene (Malek, Lättig, Hiesel, Brennicke, & Knoop, 1996) in line with an extended taxon sampling of more land plant clades. The mitochondrial small subunit ribosomal 19S rRNA (Duff & Nickrent, 1999) was likewise established to investigate a phylogeny sampling all major land plant groups. The mitochondrial nad5 gene was introduced into early land plant phylogeny as a further protein encoding locus both with a taxon sampling focussing on pteridophytes (Vangerow, Teerkorn, & Knoop, 1999) and on

bryophytes (Beckert et al., 1999) and these studies are currently being extended to nad2, nad4 and nad7 (Fig. 1).

4. INTRONS IN MITOCHONDRIAL DNA OF BRYOPHYTES

The work on nad5 in both divisions of cryptogamous plants, bryophytes and pteridophytes, led the way to observations confirmed afterwards in other genes – the high overall positional stability of mitochondrial introns defining higher-order clades. An intron unknown to seed plants and bryophytes was identified as a characteristic for the pteridophytes (Fig. 1), absent only in *Ophioglossum* and *Equisetum* but present also in the whisk fern *Psilotum* (Vangerow et al., 1999). Today we can be reasonably confident that the earliest dichotomy in the phylogeny of recent vascular plants separates lycophytes (Lycopodiopsida, lycopsids) as the sister group from the remaining tracheophytes (Euphyllophytes), in which the "Moniliformopses" (true ferns + horsetails + whisk ferns) stand in a sister-group relationship to the spermatophytes (Raubeson & Jansen, 1992; Kolukisaoglu, Marx, Wiegmann, Hanelt, & Schneider-Poetsch, 1995; Pryer et al., 2001; Karol, McCourt, Cimino, & Delwiche, 2001).

In the bryophytes on the other hand, a group I intron (here denoted as nad5i2) known from the complete mitochondrial DNA of *Marchantia polymorpha* (Oda et al., 1992) was identified in all liverworts and mosses investigated without exception (Beckert et al., 1999). However, with hindsight this mitochondrial intron conserved between mosses and liverworts turned to be a unique exception rather than the rule. The investigation of the nad5 gene has meantime been extended to encompass all intron insertion sites known in angiosperms (Fig. 1). Two further introns nad5i1 and nad5i3, present in angiosperms (Unseld, Marienfeld, Brandt, & Brennicke, 1997), are conserved in the mosses (Pruchner & Knoop, *in preparation*). Of particular interest in this context is intron nad5i3 which is conserved in the angiosperms in a disrupted genomic arrangement requiring trans-splicing (Malek & Knoop, 1998). The now identified homologue is the first cis-arranged ancestor of a trans-splicing intron in the mosses (Knoop & Pruchner, *in preparation*). In contrast, even the somewhat isolated jungermanniid liverwort *Haplomitrium* shows exactly the same intron conservation pattern, which was observed in the distantly related liverwort *Marchantia*, namely nad5i2 as the only intervening sequence. The observations of introns shared between mosses and angiosperms to the exclusion of liverworts are extended to the nad4 and nad7 genes (Pruchner, Nassal, Schindler, & Knoop, 2001). In the mosses, only three of those mitochondrial introns are found which are known to be conserved in the angiosperms (Fig. 1): nad4i1, nad7i1 and nad7i2, none of the liverwort-type. As an example the enigmatic genus *Takakia* would be placed among the mosses already based on these data alone with the typical moss-like presence of these three introns (only). For intron-containing coding regions the complete picture is already obtained for the nad2 gene in which we observe different, but strictly conserved intron occurrences within each class (Fig. 1): nad2i1 only in mosses, nad2i3 only in liverworts and nad2i3 plus nad2i4 in hornworts (Pruchner, Beckert, Muhle, & Knoop, 2002). Hence, the mitochondrial

data so far clearly support the classical systematic assessments and the definition of classes, which coincide without exception with the presence of certain introns in mitochondrial genes. Altogether, considering nad2, nad4, nad5 and nad7, six mitochondrial introns (all of group II intron type) are shared between mosses and angiosperms, only one between mosses and liverworts (nad5i2, the only group I type intron) and only one between liverworts and angiosperms (nad2i3). While this apparently places mosses in closer neighborhood to seed plants, the hornworts so far remain the major information gap. We assume that the problems to obtain complete

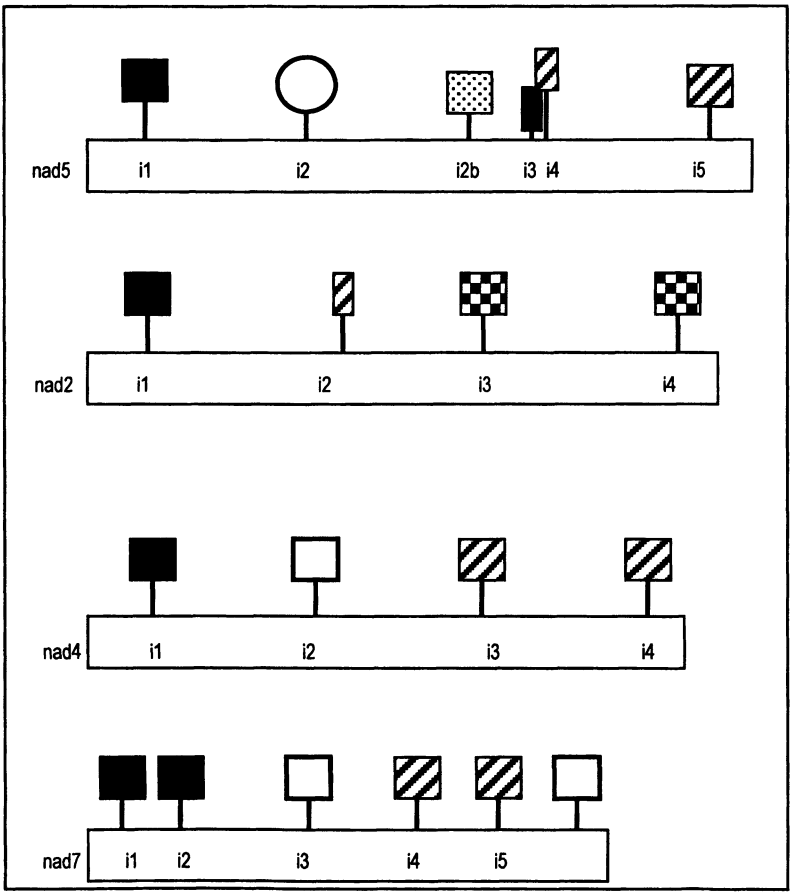

Figure 1: Structures of the mitochondrial nad genes under investigation in one of the authors' laboratory. Squares depict group II introns, slim rectangles trans-disrupted group II introns in angiosperms and a circle depicts the only group I intron in the genes under investigation, present in liverworts and mosses only. Open squares depict group II introns identified in the liverwort *Marchantia polymorpha* mitochondrial genome, hatched rectangles depict conserved angiosperm introns. Filled rectangles are angiosperm introns found conserved in the mosses. Intron nad2i3 is the only intron conserved between angiosperms and the liverwort and also present in hornworts whereas intron nad2i4 is conserved between angiosperms and hornworts only. Intron nad5i2b is a characteristic for most pteridophytes.

mitochondrial gene sequences via PCR are a problem due to the frequent and two-directional RNA editing in the hornworts (see below).

5. OBSERVATIONS IN THE CHLOROPLAST DNA

The complete nucleotide sequence of the chloroplast DNA of the hornwort *Anthoceros formosae* was recently determined in one of our laboratories (Kugita et al., 2002). *Anthoceros* has the largest chloroplast genome of land plants so far analysed, partly due to an extension of the inverted repeat (IR) region. A congruence of most genomic features (e.g. presence of genes and introns) had earlier been observed in comparison of seed plant plastomes and the one of the only other bryophyte determined, the liverwort *Marchantia*. By and large this congruence of most chloroplast genomic features can now be extended to the hornwort except for singular observations: The otherwise functional genes in other land plant plastomes, matK and rps15, have become pseudo-genes in the *Anthoceros* cp DNA. *Anthoceros* cp DNA carries a group I intron in the large ribosomal rRNA (rrn23) gene at an identical position where it is present in the distant algae *Chlorella* and *Chlamydomonas* but absent from the more closely related Coleochaetales alga *Chaetosphaeridium* (Turmel, Otis, & Lemieux, 2002) and all other land plants. A scenario of independent gains (or losses) of this group I intron is thus likely. Two

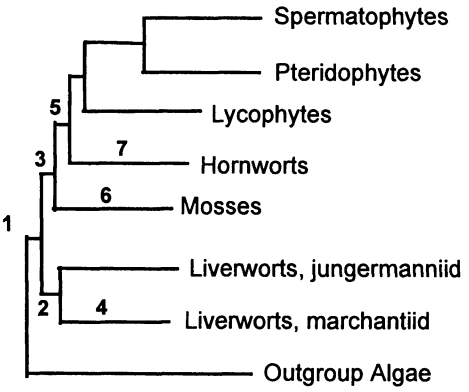

Figure 2: Characters of organellar genome evolution mapped onto a model land plant phylogeny, supported by multi-gene analyses (Qiu et al., Knoop et al., *unpublished observations*) and concatenated chloroplast protein sequences (Kugita et al., 2002). Note that significant statistical support for a sister group relationship to tracheophytes and information on genomic features in nad4 and nad7 are lacking for hornworts. 1: Gain of RNA editing, gain of nad5i2 & nad2i3. 2: Gain of many liverwort-type mitochondrial introns including nad4i2, nad7i3 and nad7i6. 3: Gain of spermatophyte-type mitochondrial introns (Qiu et al., 1998), including nad5i1, nad5i3, nad2i1, nad4i1, nad7i1 and nad7i2 and of chloroplast intron ycf3i2. 4: Loss of RNA editing activity. 5: Gain of nad5i4, loss of ycf66 and nad5i2, shift to frequent and bi-directional RNA editing. 6: Loss of nad2i3. 7: Loss of nad2i1.

introns in the ycf3 reading frame are present in *Anthoceros* in homology to seed plants, but this is also observed in the moss *Physcomitrella* (Sugita et al., *personal communication*). Like the intron in rrn23, the ycf3 intron occurrence does not give a conclusive picture of evolution, because intron 1 is present in *Marchantia* but intron 2 in the alga *Chaetosphaeridium* as single introns each. The only cladistically informative genomic feature is absence of reading frame ycf66, supporting branch 5 in figure 2, see below. Phylogenetic analysis of more than 10,000 concatenated chloroplast DNA amino acid positions results in a phylogenetic tree supporting the phylogeny shown (Fig. 2), albeit only with weak statistical confidence (Kugita et al., 2002).

6. THE BRYOPHYTE CLADES, PHYLOGENY OF BRYOPHYTES.

So far, the majority of molecular data support the grouping of extant bryophyte species in the three classically distinguished classes: hornworts, mosses and liverworts. The monophyly of liverworts was questioned by analyses of molecular 18S rRNA data sets based on which the authors suggested a sister group relationship of jungermanniid liverworts and mosses (Bopp et al., 1995; Capesius, 1995; Bopp & Capesius, 1996; Capesius et al., 1997). However, these findings were not confirmed by other analyses and could not be verified when taxon sampling was altered. Nevertheless, it should be kept in mind, that one genomic feature that could well be explained by the Bopp-Capesius topology is the so far apparent absence of organellar RNA editing in the marchantiid liverworts only (see below).

No recent molecular study has suggested the division bryophytes to be a monophyletic clade. Probably only one of the three classes or, alternatively, two classes in a joint clade share a common ancestor with vascular plants (Donoghue, 1994). Assuming one bryophyte class as sister group to all other embryophytes then still leaves us with nine topological alternatives of land plant phylogeny. Ultimately clear data demonstrating, which topological alternative would be correct, are lacking. Seeing liverworts as the basal-most land plant clade, sister to all other embryophytes, has probably so far received the strongest support from molecular data (Qiu et al., 1998) but then again, alternative views, favouring hornworts as the sister clade to all other land plants have been published both based on molecular (Nickrent, Parkinson, Palmer, & Duff, 2000; Hedderson, Chapman, & Cox, 1998) and combined molecular-morphological (Renzaglia, Duff, Nickrent, & Garbary, 2000) analyses.

Nevertheless, we here present a model land plant phylogeny (Fig. 2) that we find most suited to map the genomic features discussed above. Multi-gene analyses support this topology but sufficiently high statistical support is so far lacking (Qiu, Knoop, *unpublished observations*; (Kugita et al., 2002).

It is likely that preference of a liverwort-basal or hornwort-basal topology can be the result of long-branch attraction phenomena, inherent to the rooting problem of land plant phylogeny, given the fact that even the most closely related

extant algae to land plants (most likely of the order Charales) are only distantly related to their common ancestor with land plants.

7. THE MOSS PHYLOGENY

The occurrence of different introns posits somewhat of a problem to use appropriate outgroup sequences to root a bryophyte clade. On the other hand the faster evolving intron sequences open up more variability for within-clade resolution. This is illustrated by our investigation of a common taxon set of 42 mosses which results in congruent trees from derived from nad2 and nad5 and increased resolution upon fusion of the data sets (Beckert, Muhle, Pruchner, & Knoop, 2001). Expectedly, the mitochondrial data are not suited to address phylogenetic questions within recently radiating families or orders, such as the Hypnales, but can well reject other order or family definitions. The observations nicely fit into revised views of peristome development in Bryopsids (Vitt, Goffinet, & Hedderson, 1998) and corroborate new models of moss evolution with certain adjustments. Gradual, i.e. serial, sister group relationships of Sphagnum/Takakia, Andreaeales, Polytrichales, Diphysciales, Buxbaumiales are clear, ultimately resulting in a "crown group" dichotomy of derived genera encompassing haplolepidous vs. diplolepidous peristomes potentially placed in superorders "Dicrananae" and "Hypnanae". Comparing data sets is not fully possible, but it seems that the fusion of two mitochondrial loci, nad2 and nad5 may have more resolution to offer in a taxon sampling of 42 mosses than a fusion of one nuclear (18S rRNA) and three chloroplast (trnL-F, rps4, rbcL) loci (Newton et al., 2000) for 33 moss taxa. Most importantly, however, the two studies neither contradict each other nor other studies of reasonably wide taxon-sampling among mosses (Samigullin et al., 1998; Chiang & Schaal, 2000) for the statistically supported branches. Within the moss phylogeny, an interesting issue remaining to be resolved is the branching pattern of Funariales, *Encalypta* and *Timmia*, currently placed at the base of the "Dicrananae". This issue becomes all the more attractive given that the model plant *Physcomitrella,* as a species of the Funariales, may occupy a phylogenetically interesting position reflecting an evolutionary transition of peristome development patterns. We are confident to resolve this issue with a focused taxon sampling by fusing the forthcoming data sets of nad4 and nad7 to the already available ones (Pruchner & Knoop, *in preparation*).

8. COMPLICATING PHENOMENA IN ORGANELLAR PHYLOGENIES

8.1. Gene transfer.

The endosymbiont gene transfer of mitochondrial genes to the nucleus is a process that has been active in plants also in very recent evolutionary times (Brennicke, Grohmann, Hiesel, Knoop, & Schuster, 1993). Several examples of recent gene transfer in the angiosperms are reported (Nugent & Palmer, 1991; Adams, Daley, Qiu, Whelan, & Palmer, 2000), but the process is apparently active in earlier land plant groups as well. The functional homologue of the nad7 gene, which resides as a

pseudogene with six stop codons in the liverwort *Marchantia polymorpha* mitochondrial DNA, is nuclear encoded (Kobayashi, Knoop, Fukuzawa, Brennicke, & Ohyama, 1997). Preliminary data indicate that the degeneration of nad7 into a pseudogene in the chondriome is not unique to *Marchantia* among the liverworts (Knoop & Pruchner, *unpublished*) while it appears to be a fully functional mitochondrial gene in mosses (Pruchner, Nassal, Schindler, & Knoop, 2001). In the case of nad7 mitochondrial and nuclear sequence copies are easily distinguished by presence or absence, respectively, of the characteristic organellar group II intron sequences. However, in other cases of questionable phylogenetic assignment on the basis of organellar gene sequences it may well be worthwhile to check whether a copy transferred to the nucleus has accidentally been identified.

8.2. RNA editing.

Coding sequences of protein genes and some tRNAs, rarely also introns and other non-coding sequences, are altered by RNA editing in mitochondrial and chloroplast transcripts of land plants. Frequent exchanges of cytidines to uridines reconstitute conserved codon entities or potentially improve base pairing in the structural RNAs. In some clades, the frequency of RNA editing goes to extremes such as in the lycophyte genus *Isoetes* (Malek, Lättig, Hiesel, Brennicke, & Knoop, 1996) or in the hornworts (Steinhauser, Beckert, Capesius, Malek, & Knoop, 1999; Yoshinaga, Iinuma, Masuzawa, & Uedal, 1996). Moreover, both in hornworts and in pteridophytes the general direction of pyrimidine exchanges is also reversed and frequent exchanges from U to C are observed in approximately equal amounts in addition to the "classic" C to U exchanges. This type of RNA sequence alteration frequently reconstitutes reading frames by eliminating stop codons and thus produces functional mRNAs from what may at first have looked like a pseudo-gene on the DNA level.

So far the evolutionary patterns of RNA editing appear congruent in both organelles, although the overall frequency of pyrimidine exchanges by itself is much higher in mitochondria vs. chloroplasts by a factor of about 20. The enigmatic shift to high frequency of reverse (U to C) RNA editing is simultaneously observed in the hornwort mitochondria and chloroplasts. More importantly, in both organelles RNA editing is identified in all land plant groups but not in the marchantiid liverworts and the algae. The discrepancy between the two liverwort subgroups, marchantiid and jungermanniid, can be striking: No evidence for RNA editing is present in *Marchantia* but C to U changes are postulated for each 8th codon in the genomic nad5 region of *Haplomitrium* now analysed (Knoop & Pruchner, *unpublished*).

RNA editing has a technical impact on phylogenetic research in two ways: Highly frequent exchanges make prognosis for the genomic pyrimidine status of a gene sequence a matter of guesswork and hinder a straightforward primer design to uncover new genetic loci for phylogenetic analyses. We found this most fundamental in the analysis of mitochondrial gene sequences among hornworts. Secondly, it has been questioned whether genomic or cDNA sequences should be

used in phylogenetic analysis (Bowe & dePamphilis, 1996). We believe that the mechanisms of RNA editing specificity evolution must be elucidated first to settle this point, but for practical purposes comparative analysis (Vangerow et al., 1999) so far shows that RNA editing has a low impact only on the construction of phylogenetic trees.

9. CONCLUDING REMARKS

Molecular data have largely confirmed the three-class systematics of bryophytes. Hornworts, liverworts and mosses are monophyletic clades recognized by the absolute majority of molecular data and this coincides well with the occurrence of specific mitochondrial intron sequences. Today there is no need anymore to define unique higher order clades for *Takakia*, now confidently placed among the mosses or *Haplomitrium* with a clear position among the liverworts. The framework of a phylogeny within the clades, once they are clearly defined as monophyletic, is established and only requires the straightforward accumulation of more taxa and sequences addressing the respective systematic level under investigation. Most data have so far accumulated for mosses in which the phylogenetic relationships of most orders (and the rejection of some as paraphyletic groups) are clear. Molecular data from the two mitochondrial genes, nad2 and nad5, converge without contradiction and are consistent with branchings identified with reliability in studies using alternative molecular markers. Adding the sequences of nad4 and nad7 to the mitochondrial alignment matrix for selected model species will presumably resolve remaining questions on the level of orders and finally position remaining interesting genera with confidence, including the model taxon *Physcomitrella.*

Significantly less data are published for the liverworts and this bryophyte class will require more attention and taxon sampling while phylogeny within the genus-poor hornworts may be less of an issue. Clearly the phylogenetic branching patterns of the highest order clades in relation to the algal predecessors of embryophyte evolution and to vascular plants is not fully resolved. A striking observation is the completely incongruent occurrence of mitochondrial introns in liverworts and other embryophytes. Given that mitochondrial introns in extant algae are rare (Turmel, Otis, & Lemieux, 2002) and if one doesn't wish postulate a massive loss and reinvasion by other group II introns during evolution the observations could fit in a liverwort-basal picture of land plant phylogeny where introns came from different outer sources after separation of the liverwort and non-liverwort lineages, as reflected in the model phylogeny in figure 2. These intron sources may well be related to fungal symbiotic donors present on land at the time of embryophyte emergence (Read, Duckett, Francis, Ligron, & Russell, 2000).

A final note: This article has not attempted to repeat what has exquisitely been summarized in detail elsewhere. Most notably, the interested reader is referred to the excellent multi-author volumes "Bryophyte Biology" edited by Jonathan Shaw and Bernard Goffinet (Shaw & Goffinet, 2000a) and "Bryology for the twenty-first century" (Bates, Ashton, & Duckett, 1998).

10. REFERENCES

Adams, K. L., Daley, D. O., Qiu, Y. L., Whelan, J., & Palmer, J. D. (2000). Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. Nature, 408, 354-357.

Bates, J. W., Ashton, N. W., & Duckett, J. G. (1998). Bryology for the twenty-first century. Leeds: Maney and Son, Ltd.

Beckert, S., Muhle, H., Pruchner, D., & Knoop, V. (2001). The mitochondrial nad2 gene as a novel marker locus for phylogenetic analysis of early land plants: a comparative analysis in mosses. Mol.Phylogenet.Evol., 18, 117-126.

Beckert, S., Steinhauser, S., Muhle, H., & Knoop, V. (1999). A molecular phylogeny of bryophytes based on nucleotide sequences of the mitochondrial nad5 gene. Plant Systematics and Evolution, 218, 179-192.

Boisselier-Dubayle, M. C., Lambourdiere, J., & Bischler, H. (2002). Molecular phylogenies support multiple morphological reductions in the liverwort subclass Marchantiidae (Bryophyta). Molecular Phylogenetics and Evolution, 24, 66-77.

Bopp, M. & Capesius, I. (1995). New Aspects of the Systematics of Bryophytes. Naturwissenschaften, 82, 193-194.

Bopp, M. & Capesius, I. (1996). New aspects of bryophyte taxonomy provided by a molecular approach. Botanica Acta, 109, 368-372.

Bowe, L. M. & dePamphilis, C. W. (1996). Effects of RNA editing and gene processing on phylogenetic reconstruction. Mol.Biol.Evol., 13, 1159-1166.

Bremer, K., Humphries, C. J., Mishler, B. D., & Churchill, S. P. (1987). On cladistic relationships in green plants. Taxon, 36, 339-349.

Brennicke, A., Grohmann, L., Hiesel, R., Knoop, V., & Schuster, W. (1993). The mitochondrial genome on its way to the nucleus: different stages of gene transfer in higher plants. FEBS Lett., 325, 140-145.

Buck, W. R., Goffinet, B., & Shaw, A. J. (2000). Testing morphological concepts of orders of pleurocarpous mosses (Bryophyta) using phylogenetic reconstructions based on TRNL-TRNF and RPS4 sequences. Molecular Phylogenetics and Evolution, 16, 180-198.

Capesius, I. (1995). A Molecular Phylogeny of Bryophytes Based on the Nuclear- Encoded 18S Ribosomal-Rna Genes. Journal of Plant Physiology, 146, 59-63.

Capesius, I. & Bopp, M. (1997). New classification of liverworts based on molecular and morphological data. Plant Systematics and Evolution, 207, 87-97.

Chiang, T. Y. & Schaal, B. A. (2000). Molecular evolution and phylogeny of the atpB-rbcL spacer of chloroplast DNA in the true mosses. Genome, 43, 417-426.

Cox, C. J., Goffinet, B., Newton, A. E., Shaw, A. J., & Hedderson, T. A. J. (2000). Phylogenetic relationships among the diplolepideous-alternate mosses (Bryidae) inferred from nuclear and chloroplast DNA sequences. Bryologist, 103, 224-241.

Cox, C. J. & Hedderson, T. A. J. (1999). Phylogenetic relationships among the ciliate arthrodontous mosses: evidence from chloroplast and nuclear DNA sequences. Plant Systematics and Evolution, 215, 119-139.

De Luna, E., Buck, W. R., Akiyama, H., Arikawa, T., Tsubota, H., Gonzalez, D. et al. (2000). Ordinal phylogeny within the hypnobryalean pleurocarpous mosses inferred from cladistic analyses of three chloroplast DNA sequence data sets: trnL-F, rps4, and rbcL. Bryologist, 103, 242-256.

Donoghue, M. J. (1994). Progress and Prospects in Reconstructing Plant Phylogeny. Annals of the Missouri Botanical Garden, 81, 405-418.

Duff, R. J. & Nickrent, D. L. (1999). Phylogenetic relationships of land plants using mitochondrial small-subunit rDNA sequences. American Journal of Botany, 86, 372-386.

Edwards, D. (2000). The role of mid-palaeozoic mesofossils in the detection of early bryophytes. Philos.Trans.R.Soc.Lond B Biol.Sci., 355, 733-754.

Edwards, D., Duckett, J. G., & Richardson, J. B. (1995). Hepatic Characters in the Earliest Land Plants. Nature, 374, 635-636.

Edwards, D., Wellman, C. H., & Axe, L. (1998). The fossil record of early land plants and interrelationships between primitive embryophytes: too little and too late? In J.W.Bates, N. W. Ashton, & J. G. Duckett (Eds.), Bryology for the twenty-first century (pp. 15-43). Leeds: Maney and Son, Ltd.

Goffinet, B., Bayer, R. J., & Vitt, D. H. (1998). Circumscription and phylogeny of the Orthotrichales (Bryopsida) inferred from rbcL sequence analyses. American Journal of Botany, 85, 1324-1337.

Goffinet, B., Cox, C. J., Shaw, A. J., & Hedderson, T. A. J. (2001). The bryophyta (mosses): Systematic and evolutionary inferences from an rps4 gene (cpDNA) phylogeny. Annals of Botany, 87, 191-208.

Goffinet, B. & Hax, N. P. (2001). Bibliography of 'molecular systematic' studies of Bryophytes. I. 1985-2000. Cryptogamie Bryologie, 22, 149-155.

Graham, L. E. (1996). Green algae to land plants: An evolutionary transition. Journal of Plant Research, 109, 241-251.

Graham, L. E., Cook, M. E., & Busse, J. S. (2000). The origin of plants: body plan changes contributing to a major evolutionary radiation. Proc.Natl.Acad.Sci.U.S.A, 97, 4535-4540.

Hedderson, T. A., Chapman, R. L., & Rootes, W. L. (1996). Phylogenetic relationships of bryophytes inferred from nuclear- encoded rRNA gene sequences. Plant Systematics and Evolution, 200, 213-224.

Hedderson, T. A., Cox, C. J., & Gibbings, J. G. (1999). Phylogenetic relationships of the Wardiaceae (Musci); Evidence from 18s rRNA and rps4 gene sequences. Bryologist, 102, 26-31.

Hedderson, T. A. J., Chapman, R. L., & Cox, C. J. (1998). Bryophytes and the origin and diversification of land plants: new evidence from molecules. In J.W.Bates, N. W. Ashton, & J. G. Duckett (Eds.), Bryology for the twenty-first century (pp. 65-77). Leeds: Maney & Son Ltd.

Heinrichs, J., Proschold, T., Renker, C., Groth, H., & Rycroft, D. S. (2002). Plagiochila virginica A. Evans rather than P-dubia Lindenb. & Gottsche occurs in Macaronesia; placement in sect. Contiguae Carl is supported by ITS sequences of nuclear ribosomal DNA. Plant Systematics and Evolution, 230, 221-230.

Hiesel, R., von Haeseler, A., & Brennicke, A. (1994). Plant mitochondrial nucleic acid sequences as a tool for phylogenetic analysis. Proc.Natl.Acad.Sci.U.S.A, 91, 634-638.

Hyvönen, J., Hedderson, T. A., Smith Merrill, G. L., Gibbings, J. G., & Koskinen, S. (1998). On phylogeny of the polytrichales. Bryologist, 101, 489-504.

Karol, K. G., McCourt, R. M., Cimino, M. T., & Delwiche, C. F. (2001). The closest living relatives of land plants. Science, 294, 2351-2353.

Kenrick, P. & Crane, P. R. (1997). The origin and early evolution of plants on land. Nature, 389, 33-39.

Knoop, V. & Brennicke, A. (2002). Molecular biology of the plant mitochondrion. Critical Reviews in Plant Sciences, 21, 111-126.

Kobayashi, Y., Knoop, V., Fukuzawa, H., Brennicke, A., & Ohyama, K. (1997). Interorganellar gene transfer in bryophytes: the functional nad7 gene is nuclear encoded in Marchantia polymorpha. Mol.Gen.Genet., 256, 589-592.

Kolukisaoglu, H. U., Marx, S., Wiegmann, C., Hanelt, S., & Schneider-Poetsch, H. A. (1995). Divergence of the phytochrome gene family predates angiosperm evolution and suggests that Selaginella and Equisetum arose prior to Psilotum. J.Mol.Evol., 41, 329-337.

Kranz, H. D., Miks, D., Siegler, M. L., Capesius, I., Sensen, C. W., & Huss, V. A. (1995). The origin of land plants: phylogenetic relationships among charophytes, bryophytes, and vascular plants inferred from complete small-subunit ribosomal RNA gene sequences. J.Mol.Evol., 41, 74-84.

Kugita, M., Kaneko, A., Yamamoto, Y., Takeya, Y., Matsumoto, T., & Yoshinaga, K. The complete nucleotide sequence of the hornwort (Anthoceros formosae) chloroplast genome: insight into the earliest land plants. Nucleic Acids Research, (in press).

La Farge, C., Mishler, B. D., Wheeler, J. A., Wall, D. P., Johannes, K., Schaffer, S. et al. (2000). Phylogenetic relationships within the haplolepideous mosses. Bryologist, 103, 257-276.

Lewis, L. A., Mishler, B. D., & Vilgalys, R. (1997). Phylogenetic relationships of the liverworts (Hepaticae), a basal embryophyte lineage, inferred from nucleotide sequence data of the chloroplast gene rbcL. Molecular Phylogenetics and Evolution, 7, 377-393.

Malek, O. & Knoop, V. (1998). Trans-splicing group II introns in plant mitochondria: the complete set of cis-arranged homologs in ferns, fern allies, and a hornwort. RNA., 4, 1599-1609.

Malek, O., Lättig, K., Hiesel, R., Brennicke, A., & Knoop, V. (1996). RNA editing in bryophytes and a molecular phylogeny of land plants. EMBO J., 15, 1403-1411.

Manhart, J. R. (1994). Phylogenetic analysis of green plant rbcL sequences. Mol.Phylogenet.Evol., 3, 114-127.

McCourt, R. M. (1995). Green Algal Phylogeny. Trends in Ecology & Evolution, 10, 159-163.

Miller, H. A. (1979). The Phylogeny and Distribution of the Musci. In G.C.S.Clarke & J. G. Duckett (Eds.), Bryophyte Systematics (pp. 11-39). London & New York: Academic Press.

Mishler, B. D. (1986). A Hennigian approach to bryophyte phylogeny. Journal of Bryology, 14, 71-81.

Mishler, B. D. & Churchill, S. P. (1984). A cladistic approach to the phylogeny of the "bryophytes". Brittonia, 36, 406-424.

Mishler, B. D., Lewis, L. A., Buchheim, M. A., Renzaglia, K. S., Garbary, D. J., Delwiche, C. F. et al. (1994). Phylogenetic-Relationships of the Green-Algae and Bryophytes. Annals of the Missouri Botanical Garden, 81, 451-483.

Newton, A. E., Cox, C. J., Duckett, J. G., Wheeler, J. A., Goffinet, B., Hedderson, T. A. J. et al. (2000). Evolution of the major moss lineages: Phylogenetic analyses based on multiple gene sequences and morphology. Bryologist, 103, 187-211.

Nickrent, D. L., Parkinson, C. L., Palmer, J. D., & Duff, R. J. (2000). Multigene phylogeny of land plants with special reference to bryophytes and the earliest land plants. Molecular Biology and Evolution, 17, 1885-1895.

Nishiyama, T. & Kato, M. (1999). Molecular phylogenetic analysis among bryophytes and tracheophytes based on combined data of plastid coded genes and the 18S rRNA gene. Mol.Biol.Evol., 16, 1027-1036.

Nugent, J. M. & Palmer, J. D. (1991). RNA-mediated transfer of the gene coxII from the mitochondrion to the nucleus during flowering plant evolution. Cell, 66, 473-481.

Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N. et al. (1992). Gene organization deduced from the complete sequence of liverwort Marchantia polymorpha mitochondrial DNA. A primitive form of plant mitochondrial genome. J.Mol.Biol., 223, 1-7.

Pruchner, D., Beckert, S., Muhle, H., & Knoop, V. (2002). Divergent Intron Conservation in the Mitochondrial nad2 Gene: Signatures for the Three Bryophyte Classes (Mosses, Liverworts, and Hornworts) and the Lycophytes. J.Mol.Evol., 55, 265-271.

Pruchner, D., Nassal, B., Schindler, M., & Knoop, V. (2001). Mosses share mitochondrial group II introns with flowering plants, not with liverworts. Mol.Genet.Genomics, 266, 608-613.

Pryer, K. M., Schneider, H., Smith, A. R., Cranfill, R., Wolf, P. G., Hunt, J. S. et al. (2001). Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature, 409, 618-622.

Qiu, Y. L., Cho, Y. R., Cox, J. C., & Palmer, J. D. (1998). The gain of three mitochondrial introns identifies liverworts as the earliest land plants. Nature, 394, 671-674.

Raubeson, L. A. & Jansen, R. K. (1992). Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. Science, 255, 1697-1699.

Read, D. J., Duckett, J. G., Francis, R., Ligron, R., & Russell, A. (2000). Symbiotic fungal associations in 'lower' land plants. Philos.Trans.R.Soc.Lond B Biol.Sci., 355, 815-830.

Renzaglia, K. S., Duff, R. J., Nickrent, D. L., & Garbary, D. J. (2000). Vegetative and reproductive innovations of early land plants: implications for a unified phylogeny. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 355, 769-793.

Reski, R. (1998a). Development, genetics and molecular biology of mosses. Botanica Acta, 111, 1-15.

Reski, R. (1998b). Physcomitrella and Arabidopsis: the David and Goliath of reverse genetics. Trends in Plant Science, 3, 209-210.

Rosenberg, M. S. & Kumar, S. (2001). Incomplete taxon sampling is not a problem for phylogenetic inference. Proc.Natl.Acad.Sci.U.S.A, 98, 10751-10756.

Rydin, C. & Källersjö, M. (2002). Taxon sampling and seed plant phylogeny. Cladistics, 18, 485-513.

Samigullin, T. H., Valiejo-Roman, K. M., Troitsky, A. V., Bobrova, V. K., Filin, V. R., Martin, W. et al. (1998). Sequences of rDNA internal transcribed spacers from the chloroplast DNA of 26 bryophytes: properties and phylogenetic utility. FEBS Lett., 422, 47-51.

Schuster, R. M. (1979). The Phylogeny of the Hepaticae. In G.C.S.Clarke & J. G. Duckett (Eds.), Bryophyte Systematics (pp. 41-82). London & New York: Academic Press.

Shaw, A. J. & Goffinet, B. (2000a). Bryophyte Biology. Cambridge University Press.

Shaw, A. J. & Goffinet, B. (2000b). Molecular evidence of reticulate evolution in the peatmosses (Sphagnum), including S-ehyalinum sp nov. Bryologist, 103, 357-374.

Steinhauser, S., Beckert, S., Capesius, I., Malek, O., & Knoop, V. (1999). Plant mitochondrial RNA editing. J.Mol.Evol., 48, 303-312.

Taylor, W. A. (1995). Spores in Earliest Land Plants. Nature, 373, 391-392.

Turmel, M., Otis, C., & Lemieux, C. (2002). The chloroplast and mitochondrial genome sequences of the charophyte Chaetosphaeridium globosum: Insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants. Proc.Natl.Acad.Sci.U.S.A, 99, 11275-11280.

Unseld, M., Marienfeld, J. R., Brandt, P., & Brennicke, A. (1997). The mitochondrial genome of Arabidopsis thaliana contains 57 genes in 366,924 nucleotides. Nat.Genet., 15, 57-61.

van de Peer, Y., De Baere, R., Cauwenberghis, J., & De Wachter, R. (1990). Evolution of green plants and their relationship with other photosynthetic eukaryotes as deduced from 5S ribosomal RNA sequences. Plant Systematics and Evolution, 170, 85-96.

Vanderpoorten, A., Hedenas, L., Cox, C. J., & Shaw, A. J. (2002). Phylogeny and morphological evolution of the Amblystegiaceae (Bryopsida). Molecular Phylogenetics and Evolution, 23, 1-21.

Vangerow, S., Teerkorn, T., & Knoop, V. (1999). Phylogenetic information in the mitochondrial nad5 gene of pteridophytes: RNA editing and intron sequences. Plant Biology, 1, 235-243.

Vitt, D. H., Goffinet, B., & Hedderson, T. A. (1998). The ordinal classification of the mosses: questions and answers for the 1990s. In J.W.Bates, N. W. Ashton, & J. G. Duckett (Eds.), Bryology for the Twenty-first Century (pp. 113-123). Leeds: Maney Publishing and the British Bryological Society.

Wall, D. (2002). Use of the nuclear gene glyceraldehyde 3-phosphate dehydrogenase for phylogeny reconstruction of recently diverged lineages in Mitthyridium (Musci: Calymperaceae). Mol.Phylogenet.Evol., 25, 10-26.

Wellman, C. H. & Gray, J. (2000). The microfossil record of early land plants. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 355, 717-731.

Wheeler, J. A. (2000). Molecular phylogenetic reconstructions of the marchantioid liverwort radiation. Bryologist, 103, 314-333.

Yoshinaga, K., Iinuma, H., Masuzawa, T., & Uedal, K. (1996). Extensive RNA editing of U to C in addition to C to U substitution in the rbcL transcripts of hornwort chloroplasts and the origin of RNA editing in green plants. Nucleic Acids Res., 24, 1008-1014.

Yoshinaga, K., Kubota, Y., Ishii, T., & Wada, K. (1992). Nucleotide sequence of atpB, rbcL, trnR, dedB and psaI chloroplast genes from a fern Angiopteris lygodiifolia: a possible emergence of Spermatophyta lineage before the separation of Bryophyta and Pteridophyta. Plant Mol.Biol., 18, 79-82.

GENOMES AND DATABASES

YASUKO KAMISUGI AND ANDREW C. CUMING

Centre for Plant Sciences, University of Leeds, Leeds LS2 9JT, U.K.

Abstract. The massive investment in genome sequencing in recent years has provided an ever-increasing resource of information for biologists. Comparative analysis of gene sequences reveals the evolutionary relatedness of organisms and predicts functions for hitherto unknown genes. Among the bryophytes, the moss *Physcomitrella patens* has been the focus of intensive gene discovery (EST) programmes in both the public and private sectors. Using simple bioinformatic tools combined with powerful PCR amplification techniques, we can readily identify the transcripts of genes homologous with those known to regulate selected functions in other model species and use this information to clone the corresponding moss genes to probe their function in gene targeting experiments. Elucidation of the functions of genes in an organism occupying a basal position in the plant phylogeny will enable us to gain insights into how diverse groups of plants have recruited particular subsets of genes to participate in functions specific to their nature, during the course of evolution. In this chapter we provide some examples of how the public access databases may be used to undertake such tasks, and outline the challenges that lie ahead for the global community of molecular bryologists.

1. INTRODUCTION

The fifty years that have elapsed since Watson and Crick unveiled to the world the structure of DNA, have ushered in a revolution in our understanding of the biological world. Fundamental to this understanding is the realisation that the sequences of bases in the DNA encode the sequence of amino acids that make up proteins: the macromolecular components that carry out biochemical processes in all forms of life. The relationship between the DNA and protein sequence emerged fitfully at first, the result of painstaking genetic studies in bacteria and their viruses, before the advance of chemical techniques enabled the increasingly rapid determination, first of polypeptide sequences, and then of the sequences of bases in nucleic acids - the latter following from the facility with which specific DNA sequences could be chemically purified through DNA cloning techniques.

The pace of discovery has been ever-increasing, with developments in the automation of DNA sequence determination walking hand-in-hand with advances in data processing, such that the entire sequence content of whole genomes can now be determined and stored for analysis by anyone with access to the World Wide Web.

In recent years, we have seen the publication of the complete genomes of many species of bacteria, yeast, the model animals *Caenorhabditis elegans* (C. elegans consortium, 1998) and *Drosophila melanogaster* (Adams *et al*, 2000) and the culmination of this effort, the human genome (Lander *et al,* 2000). The plant

A.J. Wood et al. (ed.),
New Frontiers in Bryology: Physiology, Molecular Biology and Functional Genomics, 17–35.

kingdom is represented by the genomes of the angiosperms *Arabidopsis thaliana* (Arabidopsis Genome Initiative, 2000) and *Oryza sativa* (Yu *et al* 2002).

MILESTONES IN MOLECULAR GENETICS

1953: Watson & Crick propose that DNA is an anti-parallel double helix.
1958: Meselson and Stahl demonstrate that DNA replication is semi-conservative
1959: Ochoa discovers RNA polymerase
1961: Brenner, Jacob & Meselson describe messenger RNA
1961: Jacob & Monod develop the operon model of gene regulation
1963: Holley determines the first nucleotide sequence is (tRNA)
1966: Nirenberg & Khorana decipher the genetic code
1972-3: Berg, Boyer & Cohen develop DNA cloning
1977: Maxam & Gilbert, and Sanger develop techniques for DNA sequence determination. Sanger publishes the first complete genome (bacteriophage □□174)
1986: Mullis develops the Polymerase Chain Reaction
1996: The first complete genome of a eukaryote is completed: *Saccharomyces cerevisiae*
1998: The complete genome of *Caenorhabditis elegans* is published
2000: The complete sequence of *Drosophila melanogaster* is published
2000: The first plant genome, *Arabidopsis thaliana* is released in draft form
2001: The first draft of the human genome is released
2003: The "finished draft" human genome is completed. The completion of the second plant genome - *Oryza sativa* - is announced.

What is notable, when we contrast the genome sequencing efforts undertaken among the animal kingdom with those being pursued for plants, is that whereas the sequencing of animal genomes encompasses a diverse range of taxa, that for plants is focussed narrowly on a small number of species whose divergence occurred no more than 50 million years ago. Only the *Arabidopsis* and rice genomes are in any stage near completion, although other sequencing efforts are being directed toward maize and soybean.

This is justified, by those engaged in these projects, by the importance of these species as a model, in the case of *Arabidopsis*, or as crops of worldwide significance. However, it sadly ignores the potential of genome analysis on a pan-phyletic scale, to yield up insights into the evolution of gene function within the plant kingdom. There are estimated to be over half a million highly diverse species of green plants on the planet. As a group, the green plants provide an unrivalled resource for the human species, through provision of materials for shelter, fuel, food, clothing and medicines, quite apart from their contribution to the maintenance of atmospheric and climatic homoeostasis. Despite this, our exploitation of plants, as food crops, is restricted to a handful of this estimated 500,000+ species. No more than 20 species provide for the staple dietary requirements of the human population and of its domestic livestock. To concentrate genomic research within this handful of species is lamentably short-sighted on the part of, if not plant scientists, then of the bodies responsible for the funding of the plant sciences worldwide, in that it

ignores the vast wealth of untapped genetic diversity within the plant kingdom: diversity that, with the advent of transgenic plant technology can now be accessed for the improvement of our existing crop species. Since the sequencing of the *Drosophila* and *Caenorhabditis* genomes can be (and has been) supported by medically-directed national bodies and charities, using the justification that the secrets learned from the genomes of these organisms will provide insights into the evolution and functioning of the human genome, which in the (relatively) long term could result in the delivery of new therapies, then such an argument should equally be used to call for a more catholic collection of plant genomes to be determined.

2. BRYOPHYTE GENOMIC RESOURCES: A FIRST STEP INTO THE UNKNOWN

The bryophytes represent the earliest group of plants to colonise terrestrial habitats. Indeed, it can be argued that the bryophytes have been instrumental in the formation of terrestrial habitats, through their ability to colonise inhospitable substrates, and through their subsequent decomposition to commence the fertilisation of these substrates through the deposition of organic matter within them. The earliest land plants must necessarily have been able to tolerate and exploit the various extremes of the virgin habitat, and their genetic adaptations to such environmental conditions remain within their genomes for us to discover and exploit. The recent past has seen substantial progress towards the exploration of bryophyte genomes. This has been spearheaded by gene discovery programmes based on the sequencing of expressed sequence tags (ESTs) from the moss *Physcomitrella patens* and (to a smaller extent) the poikilohydric moss species, *Tortula ruralis*. Significantly, these programmes of discovery have been supported not only by public sector funding, but also by substantial investment from the private sector.

3. WHY *PHYSCOMITRELLA*?

For many years, the moss *Physcomitrella patens* has been studied as a model for understanding certain aspects of plant development - in particular, those concerning cellular differentiation and growth. The properties of the moss that recommend it for such studies have been detailed in a number of reviews, and it is not our purpose to rehearse these at length (Cove, 2000; Reski, 1999). The evolutionary distance between the origins of the mosses and the very much more intensively studied angiosperms immediately recommends the use of *P. patens* as a model for comparative studies of development, enabling us to understand the origins of some adaptations and to trace the evolution of gene function within the plant kingdom. The haploid nature of the dominant gametophyte generation, enables genetic studies to utilise the immediate recognition of mutant phenotypes, and this has led to the development of the complete range of molecular genetic techniques - in particular gene isolation and genetic transformation - being applied to this species.

This pioneering work in what was a previously unfashionable species resulted in one of the most significant discoveries in plant molecular biology: namely the

ability of *P. patens* to incorporate transforming DNA containing a sequence homologous with its genome, at the cognate genetic locus through a highly accurate and efficient system of homologous recombination. This provided a tool for the accurate engineering of a plant genome unavailable to researchers using higher plant models such as *Arabidopsis thaliana*, and more than any other single discovery resulted in a reappraisal of the utility of the "moss system" as one worthy of serious study.

In order to manipulate the genome in a predictable and useful manner however, it is first necessary to have some knowledge of that genome's coding capacity. The *Physcomitrella* genome is estimated to comprise *ca.* 511 Mbp of DNA (Schween *et al*, 2003). This DNA is distributed among 27 chromosomes - a surprisingly large number, compared with its relatives. Like most plant genomes, the *Physcomitrella* genome has probably evolved through polyploidisation events, but nevertheless the gametophyte can be seen to behave genetically as a functional haploid, although doubtless there remain (as in other plants) redundant genetic functions.

4. EST RESOURCES

EST sequencing - the high-throughput, single-pass sequencing of randomly selected cDNA clones - provides a rapid and cost-effective way of determining the information content of the expressed genome (the "transcriptome") (Rudd, 2003). The most comprehensive programme of EST sequencing for *Physcomitrella* (in terms of numbers of genes sampled) has been the commercially-funded collaboration between BASF Plant Science and Prof. Ralf Reski, at the University of Freiburg. Unfortunately, due to the proprietary nature of the programme, direct access to this sequence resource is not freely available. However, applications for sequence data held in this resource are welcomed by its owners, subject to appropriate agreements being made concerning the intellectual property rights associated with the information. Notwithstanding this restriction, valuable information about the extent and nature of the sequence database has been made available to the wider plant science community. Thus it has been published (Rensing *et al.*, 2002a; 2002b) that the collection comprises 110,087 individual cDNA sequences, obtained from libraries corresponding to a full range of *Physcomitrella* cell types and developmental stages, and subsequently subjected to competitive hybridisation ("normalisation") in order to minimise the contribution made by transcripts derived from highly expressed genes. Bioinformatic analysis of the cDNA sequences resulted in the determination that these could be clustered into 33581 non-redundant sequences, which it was estimated corresponded to approximately 25,000 genes: probably near-complete coverage of the *Physcomitrella* transcriptome. Thus any researcher seeking a sequence not represented in one of the publicly-accessible collections would have a good chance of being able to obtain that sequence from the proprietary source, upon entering into the necessary collaborative agreement.

Public access collections have been generated by two nationally funded consortia. The UK Biotechnology and Biological Sciences Research Council has

supported the Leeds University (UK) - Washington University in St. Louis (USA) "PEP" (*Physcomitrella* EST Programme; see http://www.moss.leeds.ac.uk), which has deposited over 20,000 sequences directly in GenBank, and in Japan, the NIBB-RIKEN collection of *ca.* 50,000 cDNAs has also been made immediately publicly accessible (http://www.moss.nibb.ac.jp). These two programmes differed in their approach to EST generation. The PEP collection derived from cDNA sequences derived from four types of cDNA library; one from protonemal tissue grown on a standard growth medium, one from tissue grown on medium supplemented with cytokinin, an inducer of bud formation, one from gametophore tissue and one from tissue treated with the plant growth regulator abcisic acid - a mediator of environmental stress responses. The sequences obtained corresponded largely to the 5'- sequences of partial cDNAs, thus ensuring an increased representation of protein-coding sequence over 5'- and 3'-untranslated sequences. By contrast, the NIBB-RIKEN collection has been based on libraries constructed from largely full-length cDNAs (using a highly efficient mRNA "cap-trapping") method (Carninci *et al* 1996). This produced a very high proportion of clones that comprised the complete transcript sequence - particularly useful for subsequent analysis and expression studies - although the complete sequences of these clones are not necessarily represented within the database.

The two publicly available resources have been bioinformatically assembled into a non-redundant set comprising 15,883 individual transcripts: approximately 60% of the estimated *Physcomitrella* transcriptome (Nishiyama *et al.*, 2003). Together, these two collections, although incomplete, represent a formidably useful tool for gene identification and isolation.

5. OTHER GENOMIC RESOURCES

EST sequences provide partial sequence information corresponding to individual mRNA species. In most cases, obtaining such information is only the first step in the pathway of discovery leading to the understanding of the function of a particular gene or family of genes. What the cDNA sequence lacks, even in the event that a complete cDNA sequence is available, is the all-important regulatory information encoded in the genomic sequences flanking the polypeptide coding sequence. Additionally, as in other eukaryotic species, the genes of *Physcomitrella* contain intervening non-coding sequences ("introns") that may also have a regulatory role in gene expression. This information can only be obtained by direct access to the genomic sequence, and as yet, there is no *Physcomitrella* genome project providing widespread access to such information. Therefore, the genomic sequence for individual genes requires that these genes be isolated individually for sequence determination. To this end, a number of additional resources and tools are available for use. A central part of the Leeds-based PEP was the provision of genomic resources additional to the cDNA collection. These take the form of a high-titre library of genomic fragments cloned in the bacteriophage vector lambda Fix II, and a large-insert Bacterial Artificial Chromosome library constructed in the BAC vector pCLD04541. These libraries are available - as a bacteriophage suspension in the former case, and as a gridded array of colonies on a filter, in the latter case - for

screening with cDNA probes obtained as EST clones. These resources are complementary in nature. The bacteriophage library is, so far as we are able to judge, a complete representation of the genomic content of the moss: we have yet to fail to isolate a sought-for sequence from this library. The BAC library offers ease and rapidity of clone isolation, since as an ordered, gridded array, any positive hybridisation signal immediately pinpoints the clone of interest. However, we estimate the genomic representation of this library to be only approximately 50% of the *Physcomitrella* genome. A further opportunity provided by the BAC library lies in the nature of the vector itself. The replicon, pCLD04541 is a so-called "BiBAC" vector: it can be mobilised into *Agrobacterium tumefaciens*, and subsequently will transmit its cloned insert into plants by T-DNA transformation (Tao & Zhang, 1998). We foresee that this will be of utility in determining the extent to which *Physcomitrella* genes may function in higher plants, and whether orthologous relationships between , say *Arabidopsis* and *Physcomitrella* genes may be inferred through the complementation of *Arabidopsis* mutants by *Physcomitrella* genes.

6. PCR CLONING OF *PHYSCOMITRELLA* GENES

If the above resources provide one avenue to the isolation of *Physcomitrella* genomic sequences, the ubiquitous Polymerase Chain Reaction (PCR) offers an alternative route to rapid gene isolation. The PCR reaction utilises the ability of thermostable DNA polymerases to amplify any sequences that lie between a pair of defined sequences for which oligonucleotide primers can be designed. The sequence information available in the *Physcomitrella* EST databases provides such sequence information. Using PCR, (and appropriate care in designing primers and reaction conditions) any gene represented in the EST database can be amplified from genomic DNA and cloned in a few days.

Critical factors for the success of such an approach to gene cloning are as follows: (i) confidence of the accuracy of sequence information in the database. (ii) the likely positions of introns in the sequences to be amplified. (iii) knowledge of the expected size of fragment to be amplified.

6.1 Accuracy of sequence information

Inaccuracy in EST databases arises from a number of sources. These include the inherent error-prone nature of the reverse transcriptase (a polymerase that lacks a proof-reading activity), and errors in base-calling by the sequence analysis software. These possible errors can be discounted where a sequence in the database is represented by a number of overlapping clones. The NIBB "*Physcobase*" resource displays all the overlapping sequences that make up each cDNA "contig": where a number of individual clones contain the same sequence, then a precisely matching oligonucleotide primer can be designed with some confidence. If the sequence of interest is represented by only a single cDNA clone, then greater caution must be exercised. Under such circumstances it would be best to obtain the EST clone from the relevant collection and to determine the complete sequence independently (this

will in any case be necessary if the EST sequence contains only a part of the cDNA clone)

6.2 Likely intron positions

Intron positions in orthologous genes have been highly conserved in evolution. This holds true for many of the genes the sequences of which, we have compared between *Physcomitrella patens* and *Arabidopsis thaliana*. Therefore, a comparison of the *Physcomitrella* cDNA sequence with the closest *Arabidopsis* homologue (if one exists) will often indicate the likely whereabouts of a *Physcomitrella* intron. Clearly, it is important to design primers that do not span an exon-intron boundary.

6.3 Size of fragment to be amplified

A genomic sequence containing several introns will be considerably longer than its corresponding cDNA sequence. It is therefore important to gauge the likely fragment size before commencing PCR, in order to use an appropriate extension time in the PCR reaction. Again, comparison with the corresponding *Arabidopsis* sequence may be informative, in judging the number of introns likely to occur. As to their length, a good rule of thumb is to assume that *Physcomitrella* introns will be approximately 1.5-2 times longer than those in *Arabidopsis*. However, a more exact estimate may be gained by a Southern blot analysis of *Physcomitrella* genomic DNA. This should use enzymes known to cut the cDNA clone (if possible), or otherwise utilise enzymes that reproducibly yield complete digestion of *Physcomitrella* genomic DNA. Such an analysis is particularly useful for the cloning of sequences flanking the transcribed region of the gene, that are not represented in the cDNA sequence, by using an "inverse PCR" strategy.

EXAMPLE 1: PCR CLONING OF A *PHYSCOMITRELLA* HOMOLOGUE OF A CLASS I *LEA* GENE.

The following is an outline of how EST sequence data and Southern blot analysis of genomic DNA have been used to clone the complete coding sequence of a *Physcomitrella* gene and its 5'- and 3'-flanking sequences.

1. An EST clone, (GenBank Accession No. AW145397) was identified as encoding a homologue of the wheat "E_m" protein.
2. Analysis of the higher-plant members of the *Lea-1* gene family predicted a single intron in the 39th codon.
3. Resequencing the cDNA clone confirmed the accuracy of the deposited EST sequence.
4. Southern blot analysis of genomic DNA indicated the presence of a single hybridising band with each of 4 enzymes, and two bands with an enzyme (*SstI*), known to cut within the cDNA clone (see figure 1).

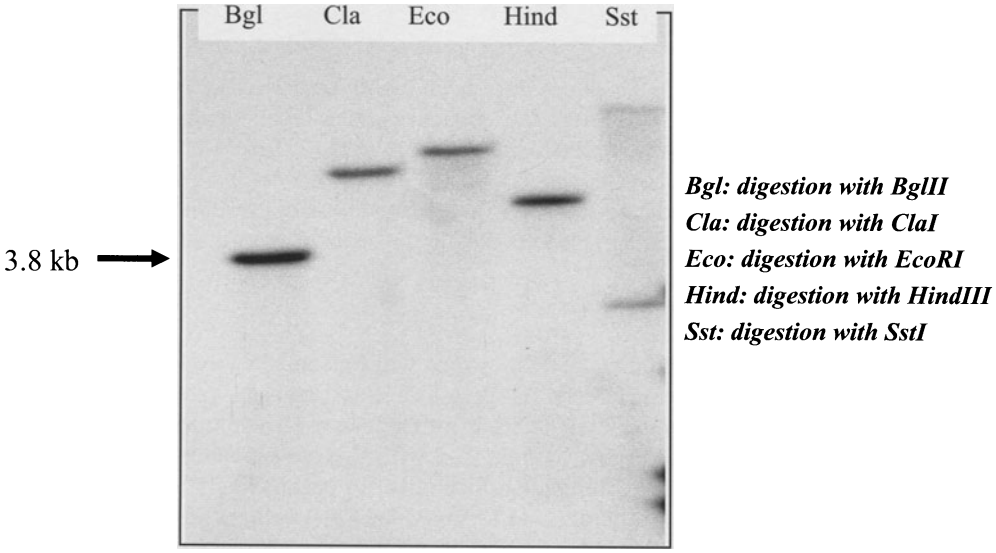

Figure 1: Southern blot of Physcomitrella DNA with PpLea-1 cDNA probe

5. Digestion with *BglII* indicated that the entire gene was carried on a single *BglII* fragment of *ca.* 3.8 kb.
6. Six PCR primers were designed (see figure 2). One pair of "inward-pointing" primers (pair 1) would be used to amplify the protein-coding sequence and its predicted intron. Two pairs of "outward-pointing" primers (pairs 2 & 3) were designed to amplify the flanking sequences by inverse PCR (iPCR, - for more details see Text Box below)

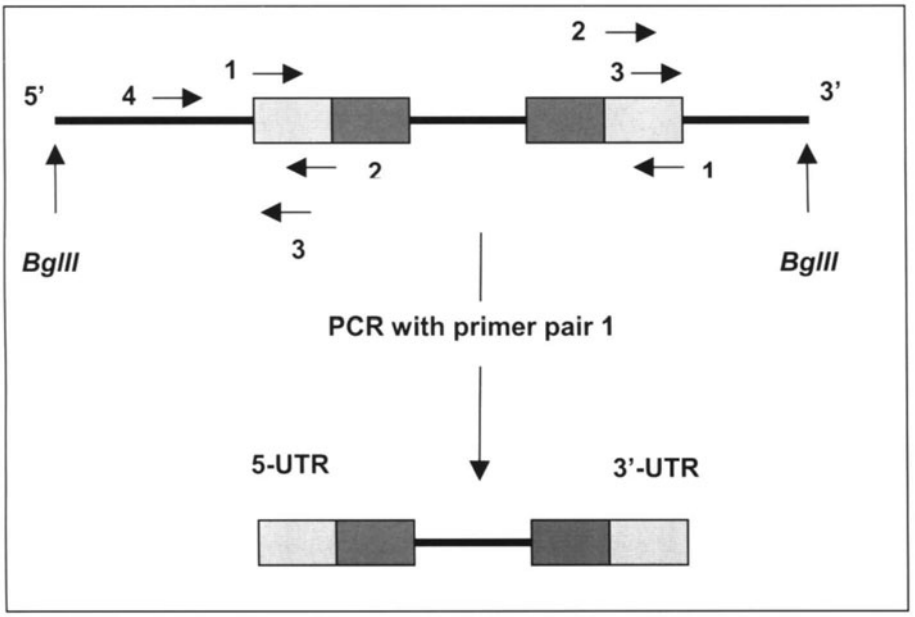

Fig 2: PCR strategy

7. The PCR with the "inward" primers (primer pair 1) amplified the genomic coding sequence. As predicted, this contained a 192 bp intron in the conserved position for this gene family.
8. Genomic DNA was digested with *BglII* and then diluted. A self-ligation reaction was performed on the diluted DNA to create predominantly circular molecules that would act as a template for inverse PCR using the pairs of "outward" primers. Two successive "nested" PCR reactions were undertaken to amplify a *ca.* 3.2kb fragment corresponding to the 5'- and 3'- flanking sequences (Figure 3). Sequence analysis of this fragment enabled the design of a seventh primer (Primer 4 – see figure 2) annealing to the presumptive 5-flanking sequence.
9. This seventh primer was used in conjunction with the 5'-end-specific "outward primers" to amplify a fragment from undigested *Physcomitrella* genomic DNA, the analysis of which confirmed that the initial iPCR product correctly corresponded to the 5'-flanking sequence of the *Physcomitrella Lea-1* homologue.

All these steps can be accomplished very rapidly: the principal source of delay is that encountered when oligonucleotides are ordered from a commercial supplier.

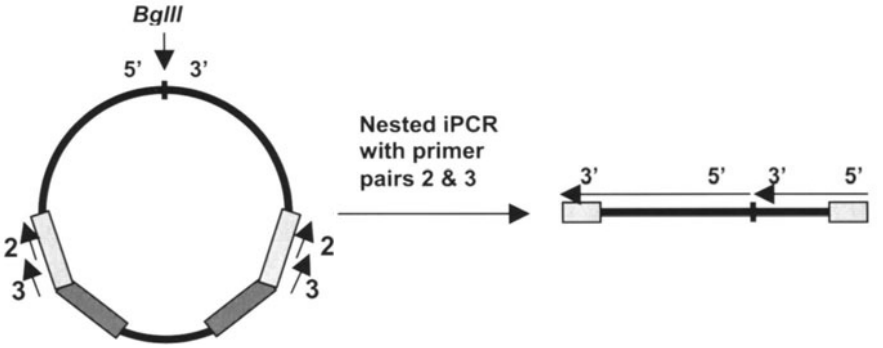

Figure 3: Inverse PCR to clone flanking sequences

INVERSE PCR (iPCR) CLONING OF GENOMIC FRAGMENTS

Step 1: Southern blotting. Southern blotting of restriction enzyme digests of genomic DNA is carried out to identify restriction fragments of suitable length for amplification by inverse PCR. In general, the aim is to identify a single fragment that hybridises with a cDNA or coding-sequence probe, in which the extent of the flanking sequences on either side of the coding sequence does not exceed a total length of *ca.* 4kb. Enzymes recommended for this purpose are those that leave overhanging complementary ends ("sticky ends") and that have been shown to digest *Physcomitrella* DNA with high efficiency. The following enzymes have proven efficacious in our hands: *BglII, EcoRI, HindIII, NcoI, NdeI* (*ClaI, HaeII, ,SstI, SstII and XhoI* yield partial digestion, but can be successfully used should the Southern blot demonstrate the reproducible production of a fragment of an appropriate length). Typically, we digest 5 µg genomic DNA which is then recovered by phenol-chloroform extraction and ethanol precipitation, prior to self-ligation.

Step 2: Self-ligation of genomic DNA. The digested DNA is dissolved in water and used to set up a ligation mixture with a total volume of 800µl. This contains 10 units of T4 DNA ligase (Invitrogen, 5units/µl). It is important to note that the

INVERSE PCR (iPCR) CLONING OF GENOMIC FRAGMENTS (continued)

reaction are 50mM Tris-Cl, pH 7.7 - 10mM $MgCl_2$ - 1mM dithiothreitol - 1mM ATP. The ligation is carried out at 16°C for at least 24 h. The reaction is heat-inactivated (65°C for 15 minutes) and concentrated to 50µl using a "Microcon30" microdialysis system (supplied by Millipore). This is then washed twice by sequential addition of 500µl 10mMTris-Cl-1mMEDTA, pH 8, each time concentrating the solution down to 50µl. The final sample can then be stored at −20°C. This provides sufficient template for 12 to 25 subsequent iPCR reactions, using different primer pairs, depending on the quantity of DNA used in each reaction.

Step 3: PCR amplification. Amplification by PCR, using "outward-pointing" primers uses standard PCR protocols. As with all PCR reactions, these must be optimised for Mg^{2+} concentration, primer annealing temperatures and extension times. The use of a proof reading thermostable polymerase is strongly recommended if the subsequent PCR product is to be cloned, prior to sequence analysis. We recommend the use of "Ex-Taq" (Takara) at a concentration of 0.25µl (1.25 units) enzyme for a 50µl reaction. A second PCR reaction using "nested" primers is frequently advisable to minimise the recovery of artefactual PCR products.

Step 4: Cloning of PCR products. There are a number of protocols and proprietary kits available for the easy cloning of PCR products. We have found the blunt-end ligation of the PCR product into the EcoRV site of the plasmid pBluescript to be both efficient and cost-effective. It is important to ensure that the vector has been effectively dephosphorylated, using calf intestinal alkaline phosphatase, in order to reduce the frequency of vector self-ligation. To maximise the efficiency with which PCR products may be cloned by blunt-end ligation, it is important to ensure that the DNA fragments are rendered flush-ended. This is achieved in a "polishing" reaction using the Klenow fragment of DNA polymerase I, which retains the 3' to 5' "proofreading" exonuclease activity.

The PCR reaction (50µl) is adjusted to reduce the initial pH of 9.0 to 9.5 by the addition of 1µl of 1M Tris-Cl pH 7.0 (Note: this is the recommended quantity if the iPCR reaction used "Ex-Taq". However, if other thermostable polymerases are used, that use buffers of different optimal pH, then the quantity of Tris pH 7 that should be added must be empirically determined: a target pH of 7.5 - 8 is recommended). 1µl of "Klenow enzyme" (7 to 10 units/µl) is added and the reaction is incubated at 25°C for 30 min to 1 h.

INVERSE PCR (iPCR) CLONING OF GENOMIC FRAGMENTS (continued)

After "polishing", the DNA is recovered by ethanol precipitation and dissolved in 20µl 10mM Tris-Cl, 1mM EDTA, pH 8.0. The PCR product is then finally purified by agarose gel electrophoresis, in order to eliminate any spurious fragments that may have been generated in the PCR amplification. The required bands are excised from a Tris-acetate buffered agarose gel and the DNA is recovered using a proprietary kit. (Our preference is for the "Gene-Clean II kit: QBiogene Inc.). Typically, the DNA is recovered in a volume of *ca.* 15 to 18µl water.

The DNA fragment is then phosphorylated in a 25µl reaction buffered by addition of 5 × DNA ligase buffer (Invitrogen) and containing mMATP and 10 units T4 polynucleotide kinase (1µl Invitrogen PNK @ 10units/µl). This is incubated at 37°C for 1 h, and the enzyme is then heat-inactivated at 65°C for 15 min. The use of the Invitrogen DNA ligase buffer to buffer this reaction enables the DNA to be used for ligation with the vector DNA without further purification. A typical ligation reaction would contain, in addition to the phosphorylated DNA fragment, 2µl dephosphorylated vector (15 to 20ng/µl), additional ligase buffer and ATP (appropriate for whatever final volume is used: for example, in a 15µl reaction containing 10µl phosphorylated DNA fragment, then a further 1µl ligase buffer would be added. A further supplementation with ATP (1µl of a 10mM stock) would be advisable insurance against exhaustion or degradation of the ATP supplied in the initial phosphorylation reaction) and 0.5µl DNA ligase (5units/µl). The volume of dephosphorylated DNA to be added will depend on the concentration of the fragment recovered from the gel, the aim being to provide equimolar quantities of vector and insert DNA. Ligation is carried out overnight at 16°C, and aliquots of the ligation product are used for transformation of *E. coli*. Recombinants are identified as white colonies on plates containing ampicillin, X-Gal and IPTG.

7. *PHYSCOMITRELLA* DATABASE SEARCHES

The publicly deposited EST sequences provide a rich resource for bioinformaticians. They also provide a useful resource for researchers who (like the authors of this chapter) are less comfortable with bioinformatic tools, but who have a rudimentary working familiarity with some of the more basic sequence analysis programs. What follows are examples of "Bioinformatics for Dummies"; some simple searches and sequence analyses that can be undertaken using sequence similarity tools such as

BLAST on websites maintained by each of the institutions that have generated EST collections, together with other easily accessible www-based resources.

The PEP site (http://www.moss.leeds.ac.uk) was the first dedicated *Physcomitrella* database, and provides a basic search capability that interrogates the individual EST sequences. More recently, a more sophisticated set of options have been provided at Freiburg (http://www.cosmoss.org) which offers options of searching the non-redundant set of contigs in addition to searches of the raw data. Currently one of the most useful databases is the NIBB "*Physcobase*" resource (http://www.moss.nibb.ac.jp) which permits searches of the non-redundant contig set, together with information on how the contigs were assembled, and links to the closest "hits" in the wider sequence databases. Additionally the "Sputnik" programme (http://mips.gsf.de/proj/sputnik/physcomitrella/) is also developing a very varied range of search tools by which the EST database may be interrogated, including potential microsatellite sequences, inverted repeat/hairpin structures, protein functional and structural features and a host of other categories. Finally, all the public EST sequences are also available within the complete sequence resources of GenBank, most commonly accessed via the National Center for Biotechnology Information (NCBI) at the US National Institutes of Health (http://www.ncbi.nlm.nih.gov). This resource has the merit of permitting keyword searches: a particularly useful way of identifying *Physcomitrella* sequences associated with particular biological functions. (This facility is limited, however, to sequences submitted by the PEP consortium, since these were annotated with the "best database hit" where such a similarity was found, prior to submission).

EXAMPLE 2: KEYWORD INITIATED SEARCHES

We shall take as an example, a function of interest to a number of *Physcomitrella* researchers who are focussing on the regulation of cellular polarity and apical tip growth. Actin depolymerizing factors have been implicated in regulating the dynamics of the actin cytoskeleton, a potentially key player in orchestrating polar growth phenomena, and are therefore candidates for functional analysis by gene knockout.

It would therefore be useful

(i) to identify any such representatives in the EST database
(ii) to determine the extent of sequence information available for any such representatives
(iii) to devise a strategy for gene isolation along the lines described in example 1, above.

Step 1: Keyword search. If the search term "Physcomitrella AND actin depolymerizing factor" is used in the NCBI database, an EST entry is returned:

```
5: BU052721 gd55h08.y1 Moss EST library PPAS Physcomitrella patens cDNA clone
PEP_SOURCE_ID:PPAS031616 5' similar to TR:Q9XEN2 Q9XEN2 ACTIN
DEPOLYMERIZING FACTOR ;, MRNA sequence|22492798|gb|BU052721.1|[22492798]
```

How reliable is this annotation? The database entries were annotated according to their similarity to other database entries, using an automatic selection of the best BLAST score. Sometimes, these annotations are spurious, the level of significance being low, or the similarity of the sequence being superceded by more recent entries with alternative functions assigned to them.

Step 2: Confirm putative identity. It is therefore wise to check this annotation by extracting the sequence and comparing it with , for example, the *Arabidopsis* sequence database. This is done simply by displaying the EST sequence in FASTA format, and using the "paste" function to enter it into the NCBI BLAST server. Since the cDNA sequence should correspond to an open reading frame, the TBLASTX function will allow the *Arabidopsis* database to be interrogated.

```
>gi|18408118|ref|NM_114470.1|   Arabidopsis thaliana actin depolymerizing
factor 1 (ADF1)(At3g46010) mRNA, complete cds Length = 835

  Score =  121 bits (259), Expect(2) = 2e-38
  Identities = 50/67 (74%), Positives = 57/67 (85%)
  Frame = -2 / +2

  Query: 531 NCQKSKIFFIAWSPSISRVRSKMIYASSKDKFKRELSGIHYELQATDPTEMDLEVIKERA 352
             NCQKSKIFFIAW P I++VRSKMIYASSKD+FKREL GI  ELQATDPTEMDL+V + RA
  Sbjct: 377 NCQKSKIFFIAWCPDIAKVRSKMIYASSKDRFKRELDGIQVELQATDPTEMDLDVFRSRA 556

  Query: 351 K*MKRSK 331
             *  ++K
  Sbjct: 557 N*KVKTK 577

  Score = 58.3 bits (121), Expect(2) = 2e-38
 Identities = 21/35 (60%), Positives = 25/35 (71%)
```

This shows a highly significant level of identity with the C-terminus of the *Arabidopsis* protein, and demonstrates that, notwithstanding the original annotation stating that this particular EST was derived from 5'-sequence information, this EST sequence actually commenced from the 3'-end. (Note: This EST clone was derived from an ABA-treated cDNA library, which, unlike the other cDNA libraries, was not constructed by directional cloning of the cDNA inserts).

Step 3: Search the *Physcobase* contigs. We can now take the EST DNA sequence and search the *Physcobase* contigs to identify the complete coding sequence. In this case a BLASTN search identifies a single, perfectly matched *Physcomitrella* contig:

```
  Sequences producing significant alignments:          (bits)  Value

  gnl|contig|Contig3315  Contig3315                    1176      0.0
```

Step 4: Obtain contig sequence and analyse. Clicking on the link to contig 3315 shows us that this contig was built from 13 overlapping clones, with a high degree of sequence reliability along its length. This enables us to design oligonucleotides for

gene amplification with a high degree of confidence. The consensus DNA sequence can be extracted and conceptually translated using a "translate" tool (*e.g.* such as http://us.expasy.org/cgi-bin/dna_aa), and the resulting output allows us to identify an open reading frame that can be seen to be highly similar with the *Arabidopsis* sequence above.

P. patens contig 3315 translation
MANASSGVAVSDDCKLKFQELQRKKAFRFIVFKIDDKVQHITVEKCGGPDATYEE
FAAALPENDCRYGVYDFDFTAEDGEINCQKSKIFFIAWSPSISRVRSKMIYASSKDK
FKRELSGIHYELQATDPTEMDLEVIKERAK

Finally a BLASTP pairwise comparison with the *Arabidopsis* sequence reveals the extent of the similarity:

```
Score =  195 bits (495), Expect = 2e-49
Identities = 97/142 (68%), Positives = 115/142 (80%), Gaps = 5/142 (3%)

Query: 1   MANASSGVAVSDDCKLKFQELQRKKAFRFIVFKIDDKVQHITVEKCGGPDATYEEFAAAL 60
           MANA+SG+AV DDCKL+F EL+ K+  RFIV+KI++K + + VEK G P  TYEEFAA L
Sbjct: 1   MANAASGMAVHDDCKLRFLELKAKRTHRFIVYKIEEKQKQVVVEKVGQPIQTYEEFAACL 60

Query: 61  PENDCRYGVYDFDF-TAEDGEINCQKSKIFFIAWSPSISRVRSKMIYASSKDKFKRELSG 119
           P ++CRY +YDFDF TAE    NCQKSKIFFIAW P I++VRSKMIYASSKD+FKREL G
Sbjct: 61  PADECRYAIYDFDFVTAE----NCQKSKIFFIAWCPDIAKVRSKMIYASSKDRFKRELDG 116

Query: 120 IHYELQATDPTEMDLEVIKERA 141
           I  ELQATDPTEMDL+V + RA
Sbjct: 117 IQVELQATDPTEMDLDVFRSRA 138
```

The databases can also be used to define the similarities and differences that exist between *Physcomitrella* genes and those of other species. In its simplest form, this may provide (i) a guide to phylogenetic relationships or (ii) an indication, where genes exist in small multigene families (*i.e.* are paralogous) as to which members of the gene families, among a number of species are most likely to be orthologous. (The definition of orthologous genes is that these fulfil essentially similar functions in different species). Where the intention is to determine gene function in *Physcomitrella* through homologous recombination-targeted gene mutation, it is clearly important to identify those genes in other species that are the orthologues of the *Physcomitrella* candidate: knowledge of this relationship will inform the search for, or analysis of the mutant phenotype.

EXAMPLE 3: WIDER SEQUENCE COMPARISONS

Using the same Actin depolymerizing factor, we can compare the predicted *Physcomitrella* ADF polypeptide sequence with those determined for the corresponding family of *Arabidopsis* genes.

Step 1: Identification of *Arabidopsis* homologues. By using the search terms "Arabidopsis AND Actin depolymerizing factor" in the NCBI Entrez database, we retrieve 76 records. These can be displayed in FASTA format and then cut and pasted into a Notepad file. The records are modified (in this case by trimming out unwanted annotation, and numbering each entry as "At1 - At76"). The *Physcomitrella* sequence can be added to this collection. This file, saved as a text file, can then be entered into the multiple sequence alignment suite of "CLUSTAL" programs.

Step 2: Preliminary clustering and redundancy reduction. By performing a multiple alignment of all these sequences (which represent a highly redundant collection, derived from EST, cDNA cloning and genomic sequencing depositions) it is easy to identify identical polypeptide sequences, derived from multiple versions of the same gene product. These can be removed from the alignment. This yields, in addition to the *Physcomitrella* sequence, a residual 21 *Arabidopsis* sequences that are used to undertake a final alignment and to produce a PHYLIP tree, based on sequence similarity (see figure 4)

8. FUTURE DEVELOPMENTS

8.1 Resources soon to be available

Recently released into the public domain is an EST collection from the desiccation tolerant moss, *Tortula ruralis*. Developed by Melvin Oliver at the USDA Agricultural Research
Station at Lubbock, Texas, this resource became available on 1 October 2003. Additionally, another means of accessing the *Physcomitrella* sequences, via gene ontology features has also been opened from this source. This latter database allows searching of the database by presumptive functional properties and can be accessed through the USDA site at http://www.csrl.ars.usda.gov. EST sequences represent merely a starting point for understanding the biology of an organism. More information is required, concerning both additional genomic sequence information, and functional information for the large numbers of genes identified by these programmes. What are the prospects for such information being forthcoming?

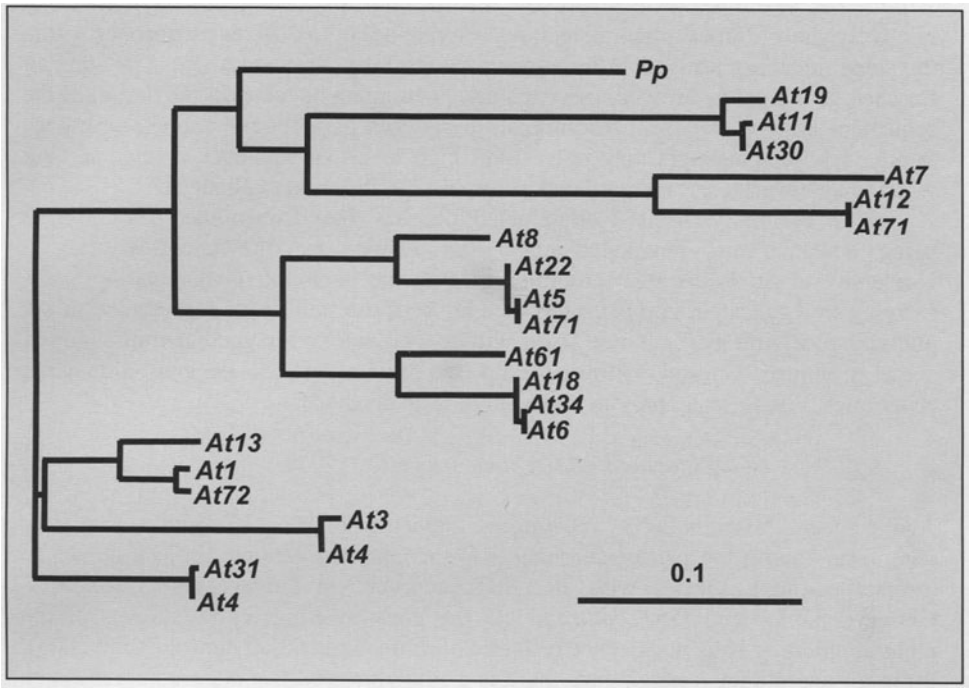

Figure 4. Sequence similarity diagram of Arabidopsis thaliana genomic sequences similar to the Physcomitrella patens putative gene specifying an actin depolymerization factor

8.2 Functional Genomics

The EST resources enable the development of tools for large-scale analysis of gene expression through the use of microarray hybridisation "gene chips". Such tools have been developed for a number of organisms, enabling patterns of gene expression associated with particular developmental, environmental or genetic challenges to be monitored on a genome-wide scale. This enables the identification of networks of co-regulated genes, many of which have hitherto undefined functions. Such arrays may be produced either by the amplification of cDNA inserts from individual EST clones that are subsequently arrayed on glass slides, or by the synthesis of oligonucleotides displayed on slides. In the case of the latter, several technologies are available, but the strategies fall into two broad classes. On the one hand, oligonucleotides can be synthesised in bulk, and robotically spotted on the glass surface using essentially the same technology that has been developed for cDNA arrays, alternatively, oligonucleotides can be synthesised, *in situ*, on the modified glass surface. The types of oligonucleotide vary according to the type of technology used. For example, in the *in situ*-synthesised chips developed by

Affymetrix, and used extensively by the *Arabidopsis* community, a number of relatively short (25mer) oligonucleotides corresponding to different sequences within the same gene, are arrayed. Alternatively, single long oligonucleotides (50-70mers) for each gene may be arrayed. In each case, care must be taken in the design of the sequences arrayed to permit discrimination between paralogous members of a gene family, whose transcripts may not exhibit high levels of sequence divergence, but which may nevertheless be regulated in response to different challenges.

An *in situ*-synthesised oligonucleotide chip (based on 60mers) is currently being designed and developed under the auspices of the Leeds-Washington University in St. Louis PEP programme, with the intention of generating slides carrying a duplicated, complete representation of the transcripts represented in the publicly accessible ESTs. These chips will be available to the general community at a cost of approximately US$400 per chip (the precise cost will depend on demand; economies of scale will apply in the event of high demand).

8.3 Will there be a Physcomitrella genome project?

At the recent "Mosses 2003" meeting of molecular bryologists, there was a clear consensus among the participants that a *Physcomitrella* genome project should be initiated. Among species with an "EST presence" in the accessible databases, *Chlamydomonas* and *Physcomitrella* are the non-flowering representatives of the plant kingdom to have a sequence resource of more than 10,000 clones (Rudd 2003). From an evolutionary standpoint, there is a pressing need to obtain comprehensive sequence information from genomes representing key positions in the land plant phylogeny, and as the costs associated with genomic sequence determination continue to fall, the cost-effectiveness of a *Physcomitrella* genome project becomes more realistic (Pryer *et al*, 2002). There are also strong strategic arguments. Approximately 6000 transcripts identified in *Physcomitrella* have no similarity with any other sequences in the various databases (Rensing *et al* 2002 b). These represent a potential reservoir of agriculturally- and pharmaceutically-useful genes about which we know nothing.

Any genome programme will necessarily be a large undertaking, and there are many gaps in our knowledge that must be filled before a large-scale sequencing effort is initiated. Most crucially, there is a gap in our understanding of the formal genetic structure of the organism. A first step in any genome-wide survey must be the development of a unified genetic and physical map of the genome to act as a skeleton upon which sequences can be anchored. Currently, a single strain of *Physcomitrella* - from Gransden Wood, near Cambridge, UK - is used in all laboratories studying this organism. Accessions from disparate geographic locations must be assembled, and tested for the extent of polymorphisms between them that can be used as genetic markers. Most likely, molecular markers such as AFLPs (Amplified Fragment Length Polymorphisms) will prove the most useful tools for the construction of a genetic map based on recombination distances, since these markers can be generated in large numbers using high-throughput, automated PCR-based techniques (Myburg & Remington: http://www4.ncsu.edu/~wagebrey/AFLP.htm). Additionally, as these markers can be directly isolated as DNA fragments, these can be integrated into a physical map by hybridisation to ordered BAC libraries.

A programme of this magnitude must also necessarily be international. Fortunately, the pool of *Physcomitrella* researchers is world-wide, with strong representation in Japan, the USA and in Europe. The challenge for this community is to mobilise a similar level of enthusiasm within their respective national funding bodies.

9. REFERENCES

Adams M.D. *et al.* (2000) The genome sequence of *Drosophila melanogaster.* Science 287, 2185-2195

Arabidopsis genome initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana.* Nature 408, 796-815

Ashton N.W. & Cove D.J. (1977) The isolation and preliminary characterisatio of auxotrophic and analogue resistant mutants of the moss *Physcomitrella patens.* Molec. Gen. Genet. 154, 87-95

C. elegans Sequencing Consortium (1998) Genome Sequence of the Nematode *C. elegans*: A Platform for Investigating Biology. Science 282, 2012-2018

Carninci P., Kvam C., Kitamura A., Ohsumi T., Okazaki Y., Itoh M., Kamiya M., Shibata K., Sasaki N., Izawa M., Muramatsu M., Hayashizaki Y. & Schneider C. (1996) High-efficiency full-length cDNA cloning by biotinylated CAP trapper. Genomics 37, 327-336

Cove D.J. (2000) The moss *Physcomitrella patens*. J. Plant Growth Regul. 19, 275-283

Lander E.S. *et al* (2001) Initial sequencing and analysis of the human genome. Nature, 409, 860-921

Myburg, A.A. & Remington, D.L. Protocol for High-Throughput AFLP Analysis Using LI-COR IR2 Automated Sequencers. http://www4.ncsu.edu/~wagebrey/AFLP.htm

Nishiyama T., Fujita Y., Shin-I .T., Seki M., Nishide H., Uchiyama I., Kamiya A., Carninci P., Hayashizaki Y., Shinozali K., Kohara Y. & Hasebe M. (2003) Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: implication for land plant evolution. Proc. Natl. Acad. Sci. USA 100, 8007-8012

Pryer K.M., Schneider H., Zimmer E.A., Banks J.A. (2002) Deciding among green plants for whole genome studies. Trends in Plant Science 7, 550-554

Reski, R. (1999) Molecular genetics of *Physcomitrella.* Planta 208: 301-309

Rensing S., Rombauts S., Van de Peer Y. & Reski R. (2002a) Moss transcriptome and beyond. Trends Plant Sci. 7, 535-538

Rensing S.A., Rombauts S., Hohe A., Lang D., Duwenig E., Rouze P., Van de Peer Y. & Reski R. (2002b): The transcriptome of the moss *Physcomitrella patens*: comparative analysis reveals a rich source of new genes.
http://www.plant-biotech.net/Rensing_et_al_transcriptome2002.pdf

Rudd S. (2003) Expressed sequence tags: alternative or complement to whole genome sequences? Trends Plant Sci. 8, 321-329

Schween G., Gorr G., Hohe A. & Reski R. (2003) Unique tissue-specific cell cycle in *Physcomitrella.* Plant Biol. 5, 1-9

Tao Q.Z. & Zhang H.B. (1998) Cloning and stable maintenance of DNA fragments over 300 kb in Escherichia coli with conventional plasmid-based vectors. Nucleic Acids Res. 26, 4901-4909

Yu J. *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp *indica*). Science 296, 79-92

PRINCIPLES OF TARGETED MUTAGENESIS IN THE MOSS *PHYSCOMITRELLA PATENS.*

DIDIER SCHAEFER AND JEAN-PIERRE ZRŸD

Laboratory of Plant Cell Genetics, Department of Plant Molecular Biology, University of Lausanne, CH 1015 Lausanne, Switzerland

Abstract: Highly efficient gene targeting is a unique feature of the moss *Physcomitrella patens*. It is now possible, in a multicellular eukaryote, to use the gene replacement technology which has been so successful in yeast. Careful design of transforming vector is nevertheless a key to successful generation of targeted plants.

1. INTRODUCTION

Functional genomics is a major field of modern biology that aims to understand the relationship between biological systems and the presence and activity of genes; model systems such as yeast have provided the biological context to explore this relationship. The recent deciphering of whole genomes from an increasing number of organisms provides a complete description of the genetic information available to fulfil their biological functions. Two major strategies are conducted in parallel to unravel the function of genes and genetic networks: (1) high throughput computational-assisted global analyses of gene expression (transcriptomics, proteomics, phenomics, reviewed in (Holtorf et al., 2002) and (2) the generation and fine characterisation of mutants (Bouche and Bouchez, 2001). The former provides an extremely valuable description of whole genome activity in response to determined experimental conditions whereas the latter enables a detailed physiological and genetic characterisation of the function of specific genes. Combining the data obtained with both approaches will provide a more accurate understanding of the function of genes and genetic networks.

Over the last decades, transgene-mediated insertional mutagenesis has replaced traditional chemical- or radiation- induced mutagenesis since subsequent identification of the mutated gene is facilitated by the presence of the inserted tag (Jeon et al., 2000; Parinov and Sundaresan, 2000). The scientific knowledge gathered from the characterisation of tagged mutations identified in the large collections of insertional mutants available in the model plant *Arabidopsis thaliana* is meaningful in that view (Bouche and Bouchez, 2001). However insertional mutagenesis by stochastic transgenesis has several methodological limitations (for a discussion see Schaefer, 2002) which can be circumvented by the capacity to target transforming sequences to specific

A.J. Wood et al. (ed.),
New Frontiers in Bryology: Physiology, Molecular Biology and Functional Genomics, 37–49.

genomic locations, i.e. to achieve efficient gene targeting (GT). The principle of GT is based on the ability to address a transforming DNA carrying a homologous targeting sequence to its corresponding chromosomal locus where it integrates by homologous recombination (HR). Gene conversion events induced in this way enable the generation of virtually any type of mutation ranging from the generation of null or weak alleles, to the fusion of molecular markers that allow the spatio-temporal expression of a gene to be followed in vivo, or by the addition of a molecular tags to the proteins to facilitate subsequent biochemical purification and characterisation.

GT is used routinely used for functional studies in micro-organisms such as bacteria, yeast and filamentous fungi but has so far found limited application in multicellular eukaryotes. This is because illegitimate integration of transgenes at random locations occurs at orders of magnitude more frequently than targeted integration at homologous loci. The current model proposes that the ratio of targeted to random integration upon transgenesis correlates with the major pathway of double strand break (DSB) repair used by the cells: most DSB are repaired by a non-homologous end joining pathway in plants and animal cells whereas yeast repairs DSB essentially by a homologous recombination pathway. However, the factors that determine this ratio are currently poorly understood and have been discussed extensively in several reviews (Mengiste and Paszkowski, 1999; Paques and Haber, 1999; Puchta, 2002; Hohe and Reski, 2003); this topic will not be discussed here. Mouse embryonic stem (ES) cells provide the only model system in animal biology that is proficient enough for GT approaches and this accounts for its exponential development over the last fifteen years (Müller, 1999). The moss *Physcomitrella patens* is the only plant in which functional genomic studies utilizing targeted transgenesis are feasible today (Schaefer and Zryd, 1997; Schaefer, 2001). We will review here the current knowledge of the features of GT in *Physcomitrella patens* and provide a description of the experimental strategies that can be used to achieve various types of mutagenesis illustrated with examples from recently obtained data.

2. CONSIDERATIONS ON VECTOR STRUCTURE AND DESIGN

Strategies for targeted mutagenesis were initially defined in budding yeast (Rothstein, 1991) and mouse ES cells (Müller, 1999). The structure of the construct plays a critical role for the transformation outcome, and the addition of a site-specific recombination system such as Cre/lox (Sauer, 1993) within the vector enables subsequent modifications of the transgenic locus (for a recent review of Cre/lox applications in plants, see Ow, 2002). Strategies combining GT with site-specific recombination provide the ultimate genetic tool for targeted mutagenesis (Müller, 1999) and form the conceptual framework for those described in this chapter for functional studies in the moss *Physcomitrella patens*. Two types of vectors are currently used for transformation: insertion and replacement vectors. Insertion vectors carry a sole homologous targeting sequence cloned next to the selectable marker (figure 1A). Insertion occurs in the target locus by a single homologous recombination event, generating an insertional mutation flanked by 2 copies of the targeting sequence. Cre-mediated removal of sequences flanked by LoxP sites will eliminate selectable markers and vector repeats , but leave a duplicated target sequence in the genome.

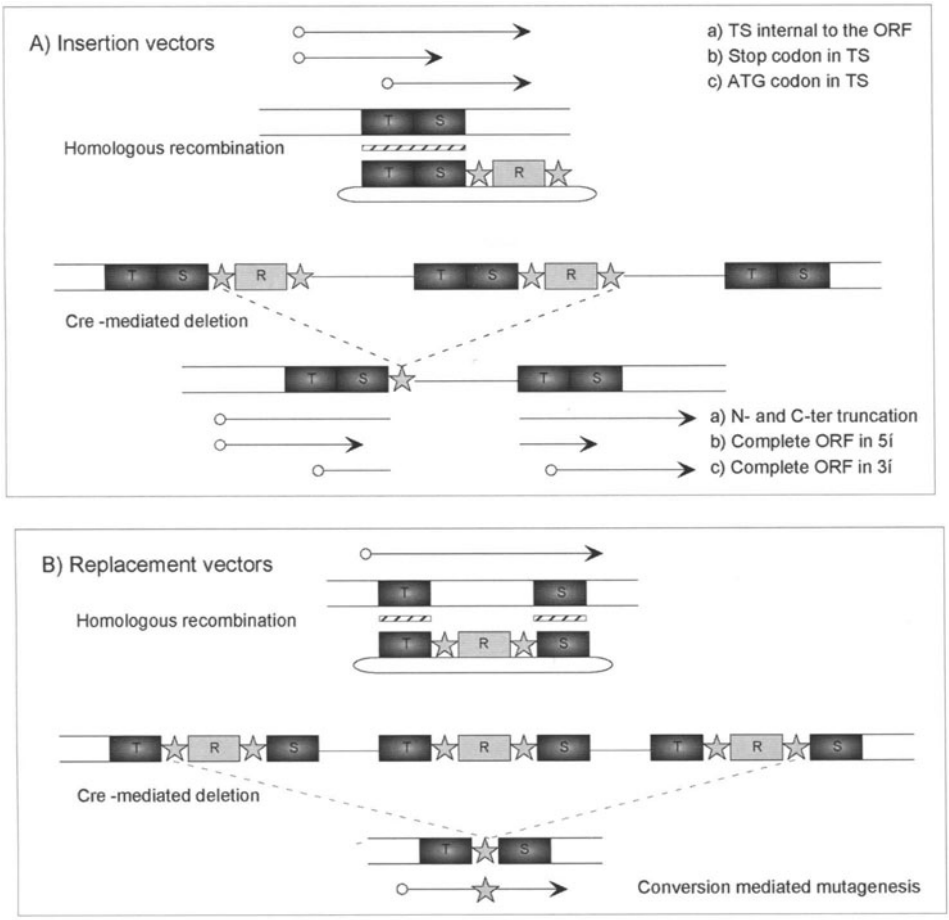

Figure 1. Basic features of insertion and replacement vectors: A) An insertion vector carries a single targeting sequence (TS boxes) cloned next to a selectable cassette (R box) flanked by 2 LoxP sites (stars). Targeted insertion is mediated by a single HR reaction (hatched bar). Cre-mediated deletions are indicated by the hatched lines. The structure of open reading frames and truncated protein after targeted insertion are illustrated with arrows, with the circle representing the ATG codon and the arrowhead the Stop codon. Note that the ATG or STOP codon must be absent from the targeting sequence since their occurrence will lead to the formation of a complete ORF with modified 5' or 3' untranslated sequences at one junction. B) A replacement vector carries the selection cassette in between 2 targeting sequence. Conversion events are mediated by 2 HR reactions and the subsequent removal of plasmid repeats leaves a single LoxP site in between targeting sequences.

Insertion vectors clearly do not provide the optimal structure for the generation of subtle mutations since gene conversion is impossible. They are also not ideal for the generation of null alleles since the remaining duplicated target sequences present in the genome

may enable the expression of truncated or mutated forms of the protein which can display transdominant properties (for a discussion of the problems observed in mouse ES cells, see Müller, 1999). The prevalence of the GT pathway within *Physcomitrella* transformation was demonstrated with insertion vectors (Schaefer and Zryd, 1997).

In replacement vectors, the selectable marker is flanked by two homologous targeting sequences, which do not have to be contiguous in the target locus but have to be in the same 5' to 3' orientation (figure 1B). Replacement vectors can integrate by two homologous recombination events and thus generate true gene conversions. Cre-mediated removal of DNA sequences flanked by LoxP sites enables the elimination of selectable markers and vector repeats, leaving a single LoxP site between the two targeting sequences, i.e. a conversion of the target sequence with the in-vitro modified one. Replacement vectors can also insert in the target locus by a single homologous recombination event taking place within one of the targeting sequence; the considerations mentioned above for insertion vectors are also valid in this case. These types of vectors have been used to assess the main parameters of GT in *Physcomitrella* and their structures provide the basis for the experimental strategies described below.

3. FEATURES OF GENE TARGETING IN *PHYSCOMITRELLA*

3.1 The history of gene targeting

The first successful transformation of Physcomitrella patens was achieved by polyethylene-glycol (PEG) mediated direct DNA-transfer into protoplasts (Schaefer et al., 1991). In the absence of sequence homology between the transforming DNA and the moss genome, relative transformation frequencies (RTF) are in average 1 in 100,000 regenerating colonies for supercoiled plasmids and ca. 5-10 times higher with linearised DNA (Schaefer, 1994). Molecular and genetic analyses demonstrate that transformation is achieved by the integration of several (2 to 30 in average) tandem repeats of the transforming plasmid usually at single location of the genome. These low RTF, compared to those observed in tobacco for example (1-0.1%), are not associated with a sub-optimal transformation protocol since transient expression of a GFP reporter reproducibly yields 5 – 25% labelled cells within 48 hours. PEG-mediated direct DNA transfer into protoplasts remains the method of choice to transform Physcomitrella.

The first evidence for the prevalence of the GT pathway in *Physcomitrella* were obtained when several moss strains transformed with a plasmid conferring hygromycin resistance were retransformed with a second plasmid conferring kanamycin resistance (Schaefer, 1994). In this situation sequence homology was provided by the plasmid backbone common to both vectors (insertion type). We observed that RTF were in average one order of magnitude higher in transformed strains than in wild type, and that the antibiotic resistance markers were genetically linked in most double resistant strains obtained by sequential transformation. Integration of tandem repeats of the second plasmid by HR with the sequences integrated in the chromosome was demonstrated at the molecular level subsequently (Schaefer and Reynolds, unpublished data). At the same time, co-segregation of resistance markers involved in sequential transformation experiments was also reported independently (Kammerer and Cove, 1996). Definitive evidence for the dominance of the GT pathway in *Physcomitrella* was provided by experiments designed to target three independent unique sequences in the moss genome (Schaefer and Zryd, 1997). This seminal work demonstrated that, in up to 90% of the

cases, insertion vectors sharing 2.3 to 3.5 kb of sequence homology with the moss genome integrate in the target locus by HR. RTF observed in these experiments was raised up to 1 in 1000 regenerating colonies and targeted transformation was associated with the insertion of direct tandem repeats of the vector. This work opened the way to the direct generation of knock-outs in *Physcomitrella* (for recent reviews see Puchta, 1998; Schaefer and Zrÿd, 2001; Puchta, 2002; Schaefer, 2002; Hohe and Reski, 2003).

3.2 The parameters influencing gene targeting

The development of GT technology required further methodological refinements to achieve gene conversions required for fine mutagenesis. The following parameters have to be assessed: (1) the relationship between the extent of sequence homology on the transforming DNA and the GT efficiency, (2) the pattern of targeted integration following transformation with replacement vectors and (3) the suitability of a site-specific recombination system such as Cre/lox to remove plasmid repeats and recycle selectable markers. In collaboration with the group of Michel Laloue (INRA Versailles), we addressed these questions in a pilot study of the adenine phosphoribosyl transferase (APRT) locus of *Physcomitrella*. This gene encodes for an enzyme of the purine salvage pathway that recycles adenine into AMP and provides a suitable target for GT studies since its loss of function confers resistance to the adenine analogue 2,6-diaminopurine (DAP). It is unique in the moss genome and is encoded at the genomic level by 7 exons with an average size of 100 bp.

The first two parameters were assessed with three different replacement vectors. In the first, a positive selectable marker was inserted in the middle of the ca. 900 bp APRT cDNA. This vector thus displayed discontinuous sequence homology with the target locus. The two other vectors independently carried 1.2 or 2.9 kb of aprt genomic sequence, in the middle of which part of the aprt coding sequence was deleted (including a specific restriction site used for subsequent molecular analyses). The relationship between GT efficiencies and the extent of sequence homology could be established as follows (Figure 2a). (1) GT is possible with cDNA based replacement vectors but it is not very efficient. Nevertheless HR events occurring within stretches of 51 and 187 bp of continuous homology have been observed. These values probably represent the lowest amount of sequence homology required for GT in *Physcomitrella*. (2) RTF (around 1 in 1000 regenerating colonies) and targeting efficiencies (95 %) were high and similar with both genomic DNA based vectors indicating that 2 stretches of 600 bp or more are sufficient to saturate GT in *Physcomitrella*. (3) Linearisation of replacement vectors at the edges of the targeting sequences is required for optimal targeting frequencies. These values are comparable to those reported in recent independent gene disruption studies (Hohe and Reski, 2003). Such a targeting efficiency can only be compared with that observed in budding yeast and signifies that *Physcomitrella* ranks above all others for the efficiency of GT among multicellular eukaryotes (Schaefer, 2001).

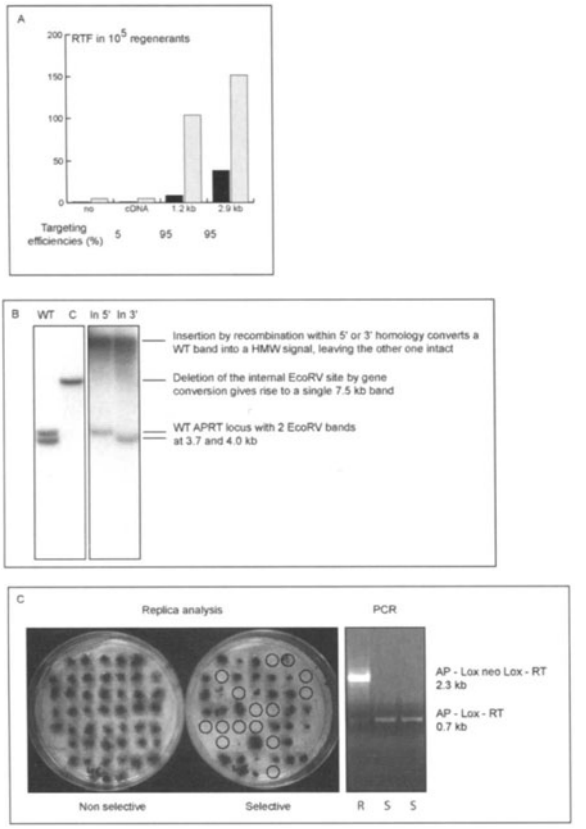

Figure 2 Parameter of GT in Physcomitrella A) Transformation frequencies (RTF) and targeting efficiencies observed with different replacement vectors on the APRT locus of Physcomitrella. RTF observed with supercoiled and linear plasmids correspond to the black and dotted bars, respectively. B) Southern blot analysis of DNA isolated from representative strains obtained following conversion (C) or insertion (In 3' and In 5') of the replacement vector in the APRT locus. DNA was digested with EcoRV and hybridised with a complete aprt sequence. The 2 WT bands (3.7 and 4.0 kb) are converted into a single band (7.5 kb) upon gene conversion, demonstrating that the deletion of 200 bp generated in the transforming DNA and containing the internal EcoRV site was effectively transferred into the genome. Only one of the WT band is shifted to high molecular weight upon targeted insertion demonstrating that targeted insertion was mediated by a single HR events occurring within one of the 2 WT bands. The HMW signal observed here corresponds to plasmid repeats that do not contain Eco RV restriction sites. C) Typical data set of the analysis of single protoplast derived colonies after transient expression of the Cre recombinase. The strains that have lost antibiotic resistance (S) upon Cre expression are circled on the selective plates. PCR analyses reveal that the resistant strains (R) carry the 2.3 kb band characteristic of the AP-Lox neo Lox–RT fragment whereas sensitive carry a new fragment of 0.7 kb which corresponds to the AP-Lox-RT fragment.

A detailed molecular analysis of the aforementioned strains revealed a strikingly predictable pattern of DNA integration. Insertion in (1 HR) and conversion of (2 HR) the aprt locus were observed at similar frequencies with the three replacement vectors tested. The pattern of insertion events was characterised by the presence of tandem direct repeats (2-30) of the whole vector as previously observed with insertion vectors. Replacement of the resident locus with a single copy of the replacement cassette accounted for 10-20% of the conversion events, and we have been able to generate specific deletions of 150 and 1100 bp in the aprt locus in this manner. The presence of tandem direct repeats of the whole plasmid and/or of the replacement cassette only was observed in the other cases (Figure 2b). The 1:1 ratio of insertion versus conversion events observed with replacement vectors was confirmed in a recent study conducted at the University of Leeds (Kamisugi et al., pers. com.) whereas the integration pattern is consistent with that described in recent independent gene disruption studies. This pattern differs from that observed in budding yeast where single copy replacement is the major class of events, but resemble that previously described for GT events observed in plants (Puchta, 2002) and animal CHO cells (Adair et al., 1998).

4. GETTING CLEAN GENE REPLACEMENT

The frequent insertion of replacement vectors requires the sorting of conversion events in the transformed population. This could in principle be achieved easily using a positive-negative screening procedure that counter selects insertion events, but the lack of a suitable negative selectable marker for *Physcomitrella* does not allow for such an approach. Therefore PCR analysis must be performed with primers specific for the recombined junction to identify gene conversion events. The presence of the selectable marker and of multiple copies of the transforming DNA is another major obstacle to the generation of fine mutations. The highly predictable integration pattern enables this situation to be solved using the site-specific recombination Cre/lox system. In this system, the Cre recombinase is able to excise or invert DNA sequences flanked by two 34 bp asymmetrical LoxP sites positioned in direct or inverted orientation, respectively (Sauer, 1993). The Cre/lox system is extensively used in transgenesis to recycle selectable markers, to eliminate plasmid repeats or to generate conditional mutations using Cre expression cassettes driven by inducible promoters. Yet the possible presence of cryptic LoxP sites in genomes may lead to additional genetic rearrangements and the use of the Cre recombinase in a transient expression assay is preferable.

Among the replacement vectors evaluated with the aprt locus, one carried a neo selectable marker flanked by 2 LoxP sites, in direct orientation, cloned in to the middle of the 2.9 kb of aprt genomic sequence. Transgenic strains obtained using this construct were used to assess the efficiency of Cre-mediated site-specific recombination in *Physcomitrella* (Chakhparonian, 2001; Schaefer et al., unpublished). Transient expression of a constitutive Cre expression cassette in protoplasts of seven strains carrying 3 to 15 direct repeats of the vector was performed. Protoplasts were regenerated at low density on non-selective medium for 2 weeks and subsequently fragments of single protoplast derived colonies were replica-plated on selective and non-selective medium. Loss of antibiotic resistance was scored after an additional week of growth and the replicates growing on non selective medium were used for subsequent amplification and molecular analyses. Within this time frame, PCR and Southern blot data can be

obtained 4 and 6 weeks after transient expression, respectively. A typical example of replica plate and PCR analyses is presented in figure 2c.

For all the strains tested, phenotypic and molecular analyses provided evidence for Cre-mediated elimination of the selectable marker in 5 to 25% of single protoplast derived colonies. Resolution of multiple copies in simple insertion or conversion events was observed in ca. 25% antibiotic sensitive colonies whereas one or two copies of the plasmid without the selectable cassette could be identified in the others, suggesting that the Cre recombinase excises marker repeats sequentially. Our data also indicated that the proportion of simple integration patterns recovered after transient Cre expression was higher in strains carrying an initial lower plasmid copy number and the choice of a transformants with few integrated repeats is recommended to optimise the probability of getting clean conversions. Nevertheless these data demonstrate that Cre-mediated recycling of selectable marker and elimination of plasmid repeats is very efficient in *Physcomitrella* and that it enables the generation of clean conversions in the moss genome in a very short time scale. It also permits the generation of multiple mutations in the same strain by sequential transformation since the elimination of plasmid repeats and selectable markers remove putative target sites for the second round of transformation and allow the same selectable marker to be used several times. Therefore LoxP sites must be integrated in the design of transformation vectors for *Physcomitrella* and we have recently developed, in collaboration with Fabien Nogué (INRA, Versailles), vectors carrying multiple cloning sites and a selectable cassette flanked by LoxP sites that are freely available upon request. With these tools in hands, the most sophisticated strategies of functional genomic by targeted mutagenesis can be applied to *Physcomitrella.*

5. SEQUENCE REQUIREMENT FOR A COMPLETE FUNCTIONAL STUDY OF A SPECIFIC GENE

Major progress has been accomplished over the last years in the amount of genetic information available for *Physcomitrella.* Two publicly funded ESTs sequencing programs have provided the scientific community with a remarkable public database covering more than 100,000 ESTs representing ca. 16,000 genes (Nishiyama et al., 2003). In parallel, another database covering 100,000 ESTs was also generated in a program associating the University of Freiburg with the private company BASF Plant Sciences (Rensing et al., 2002). A detailed analysis of the function of a gene requires a complete knowledge and access to its sequence including the flanking genomic sequences. Since the whole genome of *Physcomitrella* is not yet sequenced, PCR based strategies have to be conducted, starting from the available EST information, to isolate the corresponding genomic sequences. Starting from a partial EST, 3' and 5' RACE with nested primers on cDNA libraries, mRNA or by RT-PCR enables the isolation of the corresponding full length cDNA. Since the exon intron structure of genes is frequently conserved between *Physcomitrella* and *Arabidopsis* or rice, the putative genomic structure of the moss gene can be deciphered from sequence comparison with the structure of the angiosperm's homologue. Amplification of the genomic sequence corresponding to the full length cDNA can easily be achieved by PCR on moss genomic DNA but care should be taken with primer design to avoid matches with putative exon junctions. Finally the isolation of flanking genomic sequences can be achieved by tailed

PCR or RAGE PCR (Cormack and Somssich, 1997) using nested primers based on the previously obtained genomic information. PCR technologies provide strategies that are faster than the classical screening of cDNA and genomic libraries, and enable the isolation of a complete genomic sequence within 6 to 8 weeks. The isolation of a complete genomic locus is necessary for a thorough functional genomic study of your favourite gene.

6. THE GENERATION OF MUTATIONS

A comprehensive study of gene function must entail the generation and characterisation of different types of mutants. Null alleles provide information on complete loss of function, weak alleles enable the identification of functional domains or critical amino-acids within the protein, and the expression pattern of the gene can be studied in vivo in strains carrying the natural chromosomal environment of the gene by translational fusion with suitable reporters. In this section we discuss the different GT strategies to adopt to generate these different types of mutations (Figure 3).

The best way to generate a null phenotype is to deletie of the entire coding sequence of the target gene. This strategy guarantees the absence of truncated or mutant forms of the protein that may acquire unexpected properties such as transdominant interactions. It is the strategy that has been adopted in the *Saccharomyces* Genome Deletion Project that aims to generate null alleles for every single gene identified in the The generation of weak alleles is not as simple; it is difficult to make reliable predictions on the best strategy to adopt. Modifications of the regulatory elements of a gene may alter its expression, and this has been observed in the experiments conducted on the aprt locus. In this work, we obtained both null and weak aprt alleles, and the characterisation of the weak phenotypes revealed, in each case, that it was associated with the presence of an entire aprt coding sequence flanked by a modified 3' UTS. We hypothesised that the stability of the mRNA could be impaired and thus lead to a reduced translation level accounting for the weak phenotype. yeast genome (Giaever et al., 2002). To delete the entire coding sequence of a gene in *Physcomitrella*, the replacement vector should carry a LoxP selectable cassette inserted between the 5' and the 3' flanking genomic sequences. Cre-mediated removal of selectable marker and repeats will leave a single LoxP site between the 5' and 3' flanking sequence.

We describe below two strategies to generate weak alleles: the first will lead to the formation of truncated forms of the protein and is suitable for the functional dissection of the role of different domains within a protein; the second is probably of more general application and enables the generation of point mutations (figure 3). To achieve C-terminal protein truncation, the replacement vector should carry (from the 5' to the 3' end) (a) the N-terminal gene sequence to which an in frame Stop codon has been added at the 3' end, (b) the LoxP selectable cassette and (c) the 3' UTS. The mutagenised locus will carry, after Cre recombination, the truncated gene driven by its natural regulatory elements. To achieve N-terminal truncations, the selectable marker should be flanked in 5' by the 5' flanking genomic sequences and in 3' by the C-terminal coding sequence to which an in frame ATG has been added at the 5'end. Protein truncation can also be obtained with insertion vectors; we have recently observed a weak phenotype that was associated with C-terminal truncation of a structural

cytoskeleton gene (Pegoraro and Schaefer, unpublished results). Protein truncation thus provides a convenient way to generate weak alleles.

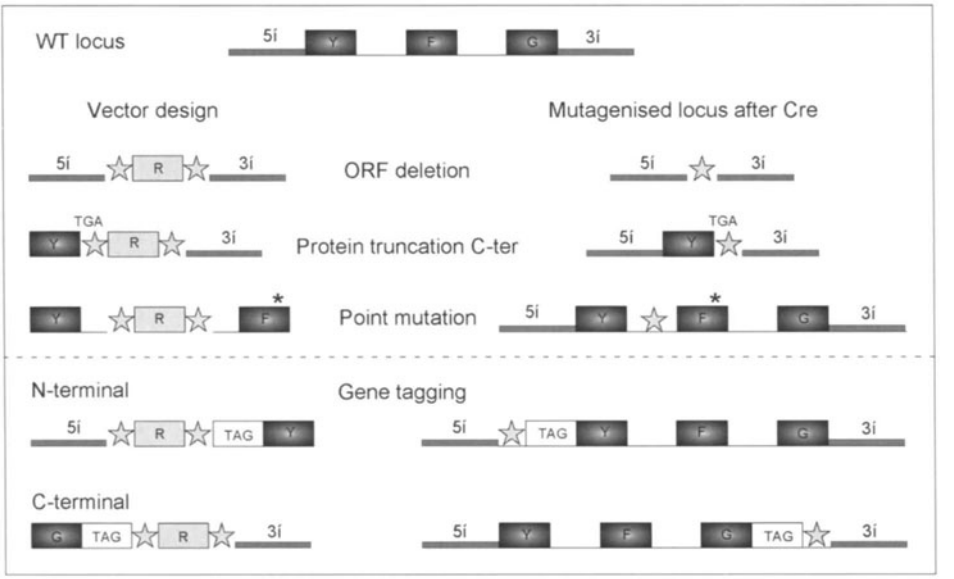

Figure 3 Schematic structure of disruption vectors designed to generate different types of mutation in the moss genome. The wild-type locus of Your Favourite Gene (YFG) is illustrated on top with 5' and 3' flanking sequences represented with thin grey bars, exons with the Y, F and G boxes and introns with thin lines. Point mutation in the F box is marked with an X on top of the box, the other symbols are identical to those used in figure 1. Please refer to the text for a detailed description of this figure.

The generation of point mutations offers the most sophisticated approach for unravelling gene function as it permits the modification of any specific amino acid within a protein. The strategy proposed here is similar to that currently used in mouse ES cells. The replacement vector carries the LoxP selectable marker cloned in an intron and flanked by the adjacent genomic sequences in which the desired mutation has been generated in vitro. The addition of a restriction fragment polymorphic trait (i.e. elimination or creation of a restriction site) to the mutated codon is recommended since it facilitates subsequent identification of the mutation. Conversion of the wild type sequence with the in vitro mutated one and elimination of selectable marker and plasmid repeats by the Cre recombinase will leave a single LoxP site in the intron. In yeast, the transfer of genetic information during HR between a transforming DNA and the chromosomal sequence is reciprocal and occurs statistically at equal frequencies. This means that ca 50 % of GT events are associated with the transfer of the mutation to the resident chromosomal information. These features are probably similar in *Physcomitrella*: this implies that the mutated strains have to be identified in the

transformed population. PCR amplification of the mutagenised sequence followed by restriction analysis or DNA sequencing allows for the identification of mutated strains. In the experiments conducted on the aprt locus, successful transfer of genetic information (in this case a deletion) from the transforming DNA to the genome was reproducibly observed providing experimental support for the feasibility of such a strategy in *Physcomitrella.*

Finally, efficient GT in *Physcomitrella* allows for the direct creation of a translational fusion of the gene under study with reporter genes (GUS, GFP; etc..) or protein tags (6 x His, TAP-tag, HA, c-Myc, etc…); this facilitates both the study of the expression profile of the protein, and its purification. The choice between N-terminal or C-terminal translational fusion cannot be predicted and this question must be addressed experimentally for each situation. The presence of an additional protein sequence to the gene may lead to unexpected side effects and a careful phenotypic analysis of the strains must be performed to validate the approach. The structure of replacement vectors designed to generate N-ter or C-ter translational fusion of moss genes with tags is depicted in figure 3. This type of approach has been successfully achieved in several studies conducted in the laboratory of M. Hasebe in Okasaki (Sakakibara et al., 2003, and personal communication).

7. CONCLUSION

The remarkable GT efficiency observed in *Physcomitrella patens* provides a unique opportunity to apply the most sophisticated genetic tool, i.e., targeted mutagenesis, to investigate the function of plant genes. The data reviewed here clearly demonstrates that it is possible to combine GT with site-specific recombination to accurately modify any sequences in the moss genome. We have described several GT strategies that can be applied in this moss for the generation of different types of mutations. This treatise is not exhaustive and the limit of such approaches probably resides within the limits of the investigator's imagination. It is a fact that *Physcomitrella* ranks above all others for the efficiency of GT in eukaryotic models and because of this the moss will be pre-eminent in the detailed study of gene function.

8. REFERENCES

Adair, G.M., Scheerer, J.B,. Brotherman, A., McConville, S., Wilson, J.H., & Nairn, R.S. (1998) Targeted recombination at the Chinese hamster APRT locus using insertion versus replacement vectors. *Somatic Cell & Molecular Genetics* 24, 91-105

Bouche, N., & Bouchez, D. (2001) *Arabidopsis* gene knockout: phenotypes wanted. *Curr Opin Plant Biol.* 4, 111-117

Chakhparonian, M. (2001) Développement d'outils de la mutagenèse ciblée par recombinaison homologue chez *Physcomitrella patens*. PhD thesis. Université de Lausanne,Switzerland "http://www.unil.ch/lpc/docs/"

Cormack, R.S., & Somssich, I.E. (1997) Rapid amplification of genomic ends (RAGE) as a simple method to clone flanking genomic DNA. *Gene* 194, 273-276

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., Arkin, A.P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtis, M., Davis, K., Deutschbauer, A., Entian, K.D., Flaherty, P., Foury, F., Garfinkel, D.J., Gerstein, M., Gotte, D., Guldener, U., Hegemann, J.H., Hempel,

S., Herman, Z., Jaramillo, D.F., Kelly, D.E., Kelly, S.L., Kotter, P., LaBonte, D., Lamb, D.C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S.L., Revuelta, J.L., Roberts, C.J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D.D., Sookhai-Mahadeo, S., Storms, R.K., Strathern, J.N., Valle, G., Voet, M., Volckaert, G., Wang, C.Y., Ward, T.R., Wilhelmy, J., Winzeler, E.A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J.D., Snyder, M., Philippsen, P., Davis, R.W., & Johnston, M. (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387-391

Hohe, A., & Reski, R. (2003) A tool for understanding homologous recombination in plants. *Plant Cell Rep.* 21, 1135-1142

Holtorf, H., Guitton, M.C., & Reski, R. (2002) Plant functional genomics. *Naturwissenschaften* 89, 235-249

Jeon, J.S., Lee, S., Jung, K.H., Jun, S.H., Jeong, D.H., Lee, J., Kim, C., Jang, S., Yang, K., Nam, J., An, K., Han, M.J., Sung, R.J., Choi, H.S., Yu, J.H., Choi, J.H., Cho, S.Y., Cha, S.S., Kim, S.I., & An, G. (2000) T-DNA insertional mutagenesis for functional genomics in rice. *Plant Journal* 22, 561-570

Kammerer, W., & Cove, D.J. (1996) Genetic analysis of the result of re-transformation of transgenic lines of the moss, *Physcomitrella patens*. *Mol. Gen. Genet.* 250, 380 - 382

Mengiste, T., & Paszkowski, J. (1999) Prospects for the precise engineering of plant genomes by homologous recombination [Review]. *Biol. Chem.* 380, 749-758

Müller, U. (1999) Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis [Review]. *Mech. Devel.* 82, 3-21

Nishiyama, T., Fujita, T., Shin, I.T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y., & Hasebe, M. (2003) Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: implication for land plant evolution. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8007-8012

Ow, D.W. (2002) Recombinase-directed plant transformation for the post genomic era. *Plant Mol. Biology* 48, 183 - 200

Paques, F., & Haber, J.E. (1999) Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae* [Review]. *Micro. & Mol. Biol. Reviews.* 63, 349-404

Parinov, S., & Sundaresan, V. (2000) Functional genomics in *Arabidopsis*: large-scale insertional mutagenesis complements the genome sequencing project [Review]. *Curr. Opin. Biotech.* 11, 157-161

Puchta, H. (1998) Towards targeted transformation in plants. *Trends in Plant Science* 3, 77-78

Puchta, H. (2002) Gene replacement by homologous recombination in plants. *Plant Mol. Biol.* 48, 173-182

Rensing, S.A., Rombauts, S., Van de Peer, Y., & Reski, R. (2002) Moss transcriptome and beyond. *Trends in Plant Science* 7, 535-538

Rothstein, R. (1991) Targeting, disruption, replacement and allele rescue: integrative DNA transformation in yeast. *Methods Enzymol.* 194, 281 - 301

Sakakibara, K., Nishiyama, T., Sumikawa, N., Kofuji, R., Murata, T., & Hasebe, M. (2003) Involvement of auxin and a homeodomain-leucine zipper I gene in rhizoid development of the moss *Physcomitrella patens*. *Development* 130, 4835-4846

Sauer, B. (1993) Manipulation of the transgene by site-specific recombination: use of cre recombinase. *Methods Enzymol.* 225, 890-900

Schaefer, D.G. (1994) Molecular genetic approaches to the biology of the moss *Physcomitrella patens*. PhD Thesis. University of Lausanne, Switzerland "http://www.unil.ch/lpc/docs/DSThesis.htm"

Schaefer, D.G. (2001) Gene targeting in *Physcomitrella patens*. *Curr. Opin. Plant Biol.* 4, 143-150

Schaefer, D.G. (2002) A new moss genetics: targeted mutagenesis in *Physcomitrella patens*. *Annu Rev Plant Biol* 53, 477-501

Schaefer, D.G., & Zryd, J-P. (1997) Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J.* 11, 1195-1206

Schaefer, D.G., & Zrÿd, J-P. (2001) The moss *Physcomitrella patens* now and then. *Plant Phys.* 127, 1430-1438

Schaefer, D.G., Zrÿd, J-P., Knight, C.D., & Cove, D.J. (1991) Stable transformation of the moss *Physcomitrella patens*. *Mol. Gen. Genet.* 226, 418-424

APPLIED GENOMICS IN *PHYSCOMITRELLA*

HAUKE HOLTORF, WOLFGANG FRANK, AND RALF RESKI

Plant Biotechnology, Albert-Ludwigs-Universität Freiburg, Schänzlestrasse 1, 79106 Freiburg, Germany

Abstract. The moss *Physcomitrella* patens, which has become a versatile tool for plant reverse genetics, has been chosen for a large scale forward genetics mutagenesis approach to address gene function at a genome wide scale. Based on the high frequency of homologous recombination found in *Physcomitrella* targeted knockout approaches of single genes have lead to new insights into the mode of action of plant genes. Besides a targeted knockout approach of around 80 target genes which are of commercial interest we have developed a transposon-mediated mutagenesis platform to mutagenize cDNA libraries derived from different developmental stages of the *Physcomitrella* life cycle. These mutagenized cDNAs are transformed into *Physcomitrella* protoplasts to generate about 80.000 transgenic *Physcomitrella* lines. The mutagenized cDNAs are supposed to integrate into the nuclear DNA at their corresponding loci thereby disrupting the underlying gene. The basis for our experimental approach was the generation of a comprehensive EST database was build from 110.000 single cDNA sequences. To realize the generation of a saturated *Physcomitrella* mutant collection we have also adapted the production of moss material and the transformation protocols to comply with the requirements of a high-throughput mutagenesis approach. The derived transgenic plants were subjected to PCR-based molecular analysis to confirm stable integration of the selectable marker cassette. Furthermore, all transgenic lines were investigated in terms of their nuclear DNA content to detect doubling of the chcromosomal DNA that may have been occurred during the transformation procedure. Each mutant plant was also subjected to a detailed analysis with respect to its physiological, morphological and metabolic traits. The frequency of altered phenotypes in the *Physcomitrella* mutant collection was considerably higher than the reported rate in Arabidopsis. Within this chapter we outline the complete procedure of the generation of the *Physcomitrella* mutant collection and their initial characterization.

1. INTRODUCTION

Recently *Physcomitrella patens* has emerged as a model system to implement plant functional genomics. *Physcomitrella* is a non-vascular, multicellular land plant and a representative of the bryophytes (Reski 1998a, Reski 1998b, Reski 1999, Schaefer 2001, Schaefer 2002). Using *Physcomitrella* as a tool for functional genomics allows one to evade some of the apparent obstacles when working with higher plant model systems (Holtorf et al. 2002a). In particular, *Physcomitrella* stands out against other systems with respect to a dominant haploid gametophytic phase and a high rate of homologous recombination.

The dominant haploid state is advantageous when performing forward genetic screens because the loss of a certain gene function is not counterbalanced by a second

A.J. Wood et al. (ed.),
New Frontiers in Bryology: Physiology, Molecular Biology and Functional Genomics, 51–70.

allele. Molecular analysis of T-DNA or transposon-mutagenised Arabidopsis knockout mutants arising from reverse or forward genetics screens have been demonstrated to result in specific gene disruptions. Only a small number of them, however, exhibit an informative phenotype that would allow the clear definition of a gene function (Bouché and Bouchez 2001). A reason for this observation is most likely the result of a high degree of genetic redundancy in Arabidopsis (The Arabidopsis Genome Initiative 2000). Precise gene disruption in *Physcomitrella* that occur in a haploid background may overcome these difficulties because single gene knockout mutants are expected to show a more comprehensive phenotype.

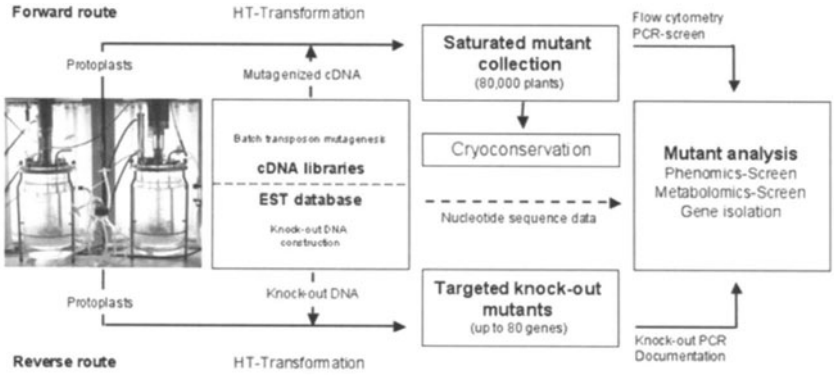

Figure 1 **Schematic representation of the high-throughput platform for production of tagged Physcomitrella patens mutants.** This scheme outlines the processes underlying the production of tagged Physcomitrella mutants. Two strategies of mutant production are illustrated: a forward and a reverse route. The forward route uses anonymous cDNA clones to produce a saturated mutant collection of 80,000 plants. The cDNAs are batch-mutagenised using the transposon shuttle mutagenesis system (see Fig.2) to disrupt the DNA coding sequence within a selection cassette. Mutagenised plasmid DNA is transferred into Physcomitrella protoplasts by way of a high-throughput transformation platform. Mutants generated by the forward route are subjected to phenotypic and metabolic profiling. The reverse route is based on known moss candidate genes, the functions of which are investigated by direct gene knockout. Disruption constructs are transfected into protoplasts and resistant plants screened by PCR for targeted transgene insertion. All mutants are subjected to further molecular analyses.

Apart from haploidy another advantage of *Physcomitrella* is even more striking. It is the only land plant studied so far that has the ability to integrate transforming DNA at a high frequency by way of gene targeting (Schaefer and Zryd, 1997), with the homologous recombination rate being several orders of magnitude higher than in seed plants (Reski 1998b, 1999). Experiments performed by several independent groups have convincingly demonstrated that a precise disruption of a genomic locus in *Physcomitrella* correlates with a mutant phenotype that clearly reveals the biological function of the targeted gene (Girke et al. 1998, Strepp et al. 1998, Girod et al. 1999, Imaizumi et al. 2002, Koprivova et al. 2002, Zank et al. 2002).

Single gene disruptions have become a routine approach in *Physcomitrella* but new strategies have also been developed to make use of the advantages of the system

for large-scale reverse genetic approaches that have the potential to complement functional gene analyses in other plant species (Reski 1998b, Schaefer 2001). In close collaboration with our partner BASF Plant Science GmbH we have started to exploit the high gene-targeting rate of *Physcomitrella* to understand plant gene function on a genomic scale. The aim is to create genetic diversity by producing a saturated mutant collection in which every expressed *Physcomitrella* gene is knocked out. It is anticipated that a high proportion of mutants will have gene lesions resulting from homologous knockout integrations. By using anonymous cDNA clones derived from all stages of moss development the entire complement of expressed moss genes was used to generate disruption constructs (Fig.1). The group of Mitsuyasu Hasebe, Okazaki, Japan used a different approach to disrupt *Physcomitrella* genes using genomic DNA instead of cDNA for the large-scale production of knockout plants (Nishiyama et al. 2000). These researchers employed a shuttle mutagenesis protocol to perform the tagged mutagenesis of genomic library material. This objective, however, cannot rule out that transgenes will be targeted to genomic non-coding regions.

Following the classical forward genetic approach mutants, that have passed selection and quality control steps, are subjected to standardized phenotypic and metabolic screening protocols. Interesting phenotypes have been described and many more are expected (Egener et al. 2000). In parallel to producing a saturated mutant collection we have also conducted approximately 80 targeted knockouts of candidate genes that are of commercial interest. These genes were selected from the EST database on the basis of homology searches against public sequence databases.

This subchapter describes the essential parts of the collaboration between the Chair of Plant Biotechnology at Freiburg University and BASF Plant Science GmbH. During the course of this project it became necessary to scale up and standardize different protocols that constitute the technological platform. The essential steps involved will be discussed in the subsequent sections.

2. MASS PRODUCTION OF STANDARDIZED MOSS MATERIAL

When intending to set up a large scale moss functional genomics project one has to realize that scaling up the process of transformation implies the development of a high-throughput platform for mass production of plant material. The well-established protocols for the transformation of *Physcomitrella* protoplasts are based on polyethylene glycol (PEG)-mediated DNA transfer (Schaefer et al.1991, Reutter et al. 1998). Early protocols for the preparation and regeneration of protoplasts from *Physcomitrella* (Stumm et al. 1975) were improved by using the enzyme mixture Driselase for efficient cell wall digestion of *Physcomitrella* protonema tissue (Grimsley et al. 1977). As a source for protoplasts this protocol uses *Physcomitrella* plants grown on solidified medium in petri dishes. The medium is supplemented with ammonium tartrate that allows cultivation of moss protonema predominantly in the chloronema stage (Ashton and Cove 1977, Jenkins and Cove 1983). Generally, chloronema cells from *Physcomitrella* are believed to be the best tissue type for the preparation of protoplasts whereas tissues from advanced developmental stages are not sufficiently susceptible to cell wall degrading enzymes (Lal 1984). The described protocols allow the preparation of sufficient amounts of protoplasts for a few transformation at a time

but are not well-suited for large scale production as the cultivation of protonema on solidified medium is very labour intensive and a constant output of transformation-competent protoplasts is not possible.

To support the need for large quantities of standardized *Physcomitrella* material bioreactor cultures are preferable. Rother et al. (1994) demonstrated that *Physcomitrella* can be cultivated in liquid mineral medium without supplying ammonium tartrate. However, under such conditions *Physcomitrella* plantlets do not preferentially grow as protonemal tissue but also develop leafy gametophores. Preparing protoplasts from a mixture of tissues at different developmental stages would be unfavourable for achieving the desired high transformation rates. Consequently, an improved bioreactor protocol was required. Other early bioreactor cultures were also unsuitable as they were grown as batch or short continuous cultures and moreover produced cell material that was not suitable for protoplast isolation (Boyd et al. 1988, Cove et al. 1997, Reutter and Reski 1996). In an attempt to generate constant amounts of suitable moss material and to reduce manual labour, semicontinuous bioreactor cultures of *Physcomitrella* were developed. Protonema tissue was pre-cultured in liquid Knop medium according to Reski and Abel (1985) and equal amounts of tissue were used to inoculate stirred tank glass bioreactors to start semicontinuous liquid cultures under varying conditions of ammonium tartrate supplementation and pH-control (Hohe and Reski 2002; Hohe et al. 2001). Manipulation of ammonium tartrate supplementation and/or pH was required in order to approximate the desired protoplast generation and transformation rates. Supplementing the growth medium with 2.5 mM ammonium tartrate led to a six fold increase in protoplast yield over normal semicontinuous *Physcomitrella* bioreactor cultures. This increase was assumed to be the result of an inhibition of caulonema development in the presence of ammonium tartrate. In contrast, rigid control of pH at 4.5 enabled higher protoplast yields despite a lack of inhibition of caulonemal development.

It was assumed that pH-conditions in the medium may influence moss cell wall composition and that a constant pH-value of 4.5 may be advantageous for subsequent enzymatic digestion of cell walls. This notion was substantiated by analysing protoplast yields from protonema grown at a constant pH value of 7.5 that led to a severe drop in numbers of isolated protoplasts. Preculture conditions which use a controlled pH value of 4.5 instead of ammonium tartrate also positively influence gene targeting (GT) efficiencies (Hohe et al. 2002). Increasing the GT efficiency through modifying culture conditions is another prerequisite for producing a saturated mutant collection that is characterized by a high rate of homologous knockout integrations.

As a result of these studies a standard semicontinuous bioreactor culture at pH 4.5, without the addition of ammonium tartrate, was used for large scale production of plant material for protoplast isolation and transformation. Semicontinuous growth conditions ensure uniformity of the starting material and avoid batch-specific variance that may disturb subsequent analysis of the moss transformants. An average harvest of around 930 ml per day from a 5 l bioreactor allows preparation of protoplasts sufficient for 105-114 transformations.

3. PRODUCTION OF A PHYSOMITRELLA EST COLLECTION

Large collections of expressed sequence tag (EST) sequences are the basis for an array of different functional genomics projects covering a wide range of different plant species (Holtorf et al. 2002a).

Such EST databases have complemented genome sequencing projects performed in both Arabidopsis and rice (The Arabidopsis Genome Initiative 2000, Yu et al. 2002, Goff et al. 2002). The combined data sets, EST and genomic sequences, allow for comparative genomic and functional genomic studies. From an evolutionary point of view bryophytes are estimated to have diverged from ferns and seed plants around 450 million years ago (Theissen et al. 2001). Comparisons between the gene content of lower and higher plants will define a set of essential genes that have been conserved during evolution. Such genes will represent the minimal complement of genes present in any plant species. Among different plant lineages divergent genes can be identified that show differences with respect to the number of gene family members, the degree of sequence conservation, and expression patterns. Early studies have already demonstrated that *Physcomitrella* is a valuable source for novel genes (Reski et al. 1998, Machuka et al. 1999).

An essential part of the functional genomics project was the generation of novel EST sequence information from three cDNA libraries specific for protonemal, gametophoric, and sporophytic material. These libraries represent almost the complete transcriptome of *Physcomitrella.* in the preparation of these cDNA libraries plant material was used from all stages of the moss life cycle and, in addition, had undergone treatments with different plant hormones. The set of cDNA libraries served two functions. First, the resulting cDNA libraries were used to perform mass sequencing of individual cDNA clones to generate a comprehensive EST database. Second, the libraries were the source of cDNAs used in the production of a saturated mutant population. Transposon mutagenesis of cDNAs, in combination with high-throughput transformation and regeneration protocols, allows targeting of every expressed *Physcomitrella* gene by way of homologous recombination. The EST collection, which mirrors the content of the underlying cDNA libraries, contains sequences of over 110,000 cDNA clones (Rensing et al. 2002a, 2002b).

Prior to mass sequencing of individual clones the protonema, gametophore, and sporophyte-specific cDNAs libraries were normalized and subtracted to drastically reduce redundancy (Rensing et al. (2002a). cDNA fragments were cloned directionally into either pUC or pBluescript plasmids and sequenced from the 3' end with primers specific to the respective vector sequences. In cases where the sequencing of a clone did not yield a valid sequence an additional sequence reaction was performed using a vector-specific 5' end primer. This was only necessary for about 3% of the sequences. Approximately 80% of the EST sequences were derived from the normalized and subtracted libraries, the remainder derive from a non-normalized protonemal library (Reski et al. 1994). In total, over 110,000 cDNA clones were sequenced representing more than five times the estimated number of *Physcomitrella* genes. This was done to ensure complete coverage of all expressed genes within the EST database. Bioinformatic tools were used for subsequent clustering of the EST sequences (Rensing et al. 2002). The removal of redundant sequences generated 33,581 sequence clusters that comprise 21,219 singletons (representing single sequences) and 12,362

contigs (built from at least two sequences). A detailed cluster analysis that generated a rate of 0.76 genes per cluster, estimated the total number of genes in *Physcomitrella* to be 25,500 +/-2,500 (expressed protein coding genes). This calculation matches the number of Arabidopsis genes, which is estimated to be 26,000, and underscores the

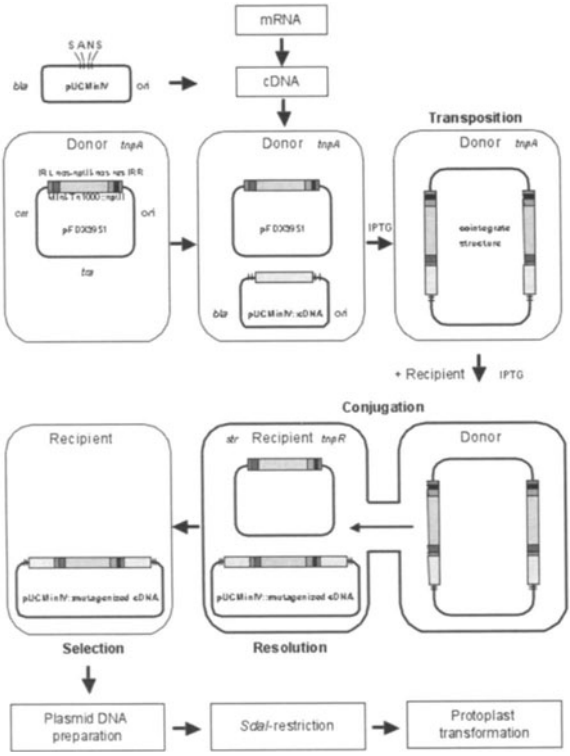

Figure 2 ***Transposon shuttle mutagenesis of Physcomitrella patens cDNA libraries.*** *Scheme outlining the process of transposon mutagenesis of Physcomitrella cDNA libraries. Low-redundancy moss cDNA populations are ligated into the minimal vector pUCMinIV. The ligation reactions are transformed into the E.coli donor strain R2217, which harbours the conjugative plasmid pFDX3951. This plasmid carries the transposon mini-Tn1000::nptII. The nptII gene is under control of nos promoter and termination signals. The donor strain also harbours the plasmid pFDX3957, which contains the bacterial transposase gene tnpA. TnpA is inducible through an IPTG-responsive promoter. IPTG induction activates expression of the bacterial transposase gene leading to transposition of mini-Tn1000::nptII. A cointegrate structure is formed between the plasmids pUCMinIV and pFDX3951. During conjugation of the donor with the recipient R1037 the cointegrate structure is transferred. Upon transfer the cointegrate is resolved by activity of the IPTG-induced resolvase gene in the recipient cell. After conjugation and resolution resultant cells are selected. Plasmid DNA is prepared from these cells and are digested with the rare-cutting restriction enzyme SdaI yielding DNA for transformation of Physcomitrella protoplasts. Plasmids are not to scale. tnpR: bacterial resolvase gene, tnpA: bacterial transposase gene, str: streptomycin resistance gene, cm: chloramphenicol resistance gene , bla: ampicillin resistance gene. Sites of restriction enzyme S (SdaI); A (AscI), N (NotI).*

notion that the moss EST database covers almost the entire transcriptome of *Physcomitrella*. While the *Physcomitrella* and Arabidopsis genomes contain a similar

number of genes, the degree of similarity among all genes is suprisingly low. Rensing et al. (2002) estimated only a maximum of 50% of the *Physcomitrella* expressed protein genes have an Arabidopsis homologue. In addition, in a comparison of the *Physcomitrella* EST database with public sequence databases using different homology search programs, 11,600 *Physcomitrella* genes failed to match any known protein sequence. Therefore *Physcomitrella* is percieved to be a rich source of unknown plant genes, a high percentage of which could be expected to exhibit novel functions. Future investigations will certainly focus on this large set of unique moss genes.

Thorough bioinformatic analyses of the EST database generated important insights into the gene content of the *Physcomitrella* genome and its homology to other plant species. In addition the EST mass sequencing approach revealed, at an early stage, that the underlying cDNA libraries are of sufficiently high quality to be used as a gene source for transposon shuttle mutagenesis (Fig.2). Cluster analysis revealed that cDNA normalization and subtraction procedures were performed successfully. This is an important criterion when the aim is to produce a *Physcomitrella* mutant population that truly saturates the genome. The low redundancy of the cDNA clones within each library helps reduce the number of independent transformations needed to disrupt every expressed moss gene.

During the course of the project, the EST database was also used for selection of candidate genes for which targeted knockout mutagenesis is desirable (see subchapter 2.4.). It also serves to validate molecular mutant analyses (see subchapter 2.8.) by providing essential sequence information regarding the identity of isolated gene sequences.

This project was the first large scale attempt to produce a comprehensive and near complete EST database for *Physcomitrella*. Other *Physcomitrella* EST sequencing projects have been launched subsequently and have generated additional ESTs. A publicly available collection of 30,000 *Physcomitrella* ESTs was generated by a collaboration between researchers of Washington University and Leeds University (http://www.moss.leeds.ac.uk). Additionally, the group of Mitsuyasu Hasebe will reslease EST sequences into the public databases. In total, around 60,000 EST sequences are currently publically available. All EST collections will be valuable sources for plant functional genomics projects and will satisfy the growing interest in the model species *Physcomitrella.*

4. *IN VIVO* TRANSPOSON MUTAGENESIS OF MOSS cDNA LIBRARIES

The ongoing moss mutant production is aimed at taking advantage of the high rates homologous recombination in *Physcomitrella* for targeted insertion of gene disruption constructs into the plant nuclear genome. The disruption constructs used are cDNA molecules that are mutagenised by the insertion of a *nptII* selection marker gene. In contrast to the approach taken by Nishiyama et al. (2000) that used genomic fragments, we used chimeric cDNAs for targeting genomic coding sequences. The use of cDNAs instead of genomic fragments as insertional transgenes should preferentially direct insertion events to expressed *Physcomitrella* genes and circumvent the possibility of a high frequency of insertions into non-coding regions. This strategy

enabled us to develop a collection of 80,000 *Physcomitrella* plants with a high percentage of targeted mutations that saturate the genome.

To allow rapid production of chimeric transgenes for every expressed *Physcomitrella* gene an efficient transposon-based shuttle mutagenesis system was developed for reliable introduction of the selectable marker gene cassette into anonymous clones from cDNA libraries (Rak et al. 2001). The system developed utilizes a bacterial transposon which has the advantage, in contrast to eukaryotic transposons, of ensuring that the transgene stays inactive inside the plant. The resulting gene disruption libraries were subsequently used to transform *Physcomitrella* protoplasts.

The process of transposon shuttle mutagenesis of *Physcomitrella* cDNA libraries is outlined in Figure 2. cDNA prepared from RNA extracted from protonemal material, grown in liquid culture for various time periods, was normalized to establish a *Physcomitrella* cDNA library representing genes expressed during vegetative growth before the onset of differentiation (Ko 1990). To reduce redundancy among cDNAs generated from gametophores and sporophytes, the reverse transcribed mRNAs were subjected to a subtraction procedure (Egener et al. 2002). The resulting low-redundancy moss cDNA populations were ligated into the minimal vector pUCMinIV, a 1.7 kb derivative of plasmid pUC19 from which most non-essential DNA sequences were deleted. The smaller vector reduced the possibility of potential transposon insertion targets within the vector sequence. The ligation reactions were transformed into the *E.coli* donor strain R2217, which harbours the novel conjugative plasmid pFDX3951. This plasmid carries the transposon mini-Tn1000::*nptII*, and retained the chloramphenicol resistance gene (*cm*) and the *tra* function from the original conjugative plasmid R388 (Avila and Cruz 1988). The mini-Tn1000 transposon, a minimalized derivative of the bacterial transposon Tn1000, is flanked by the transposon Tn1000 border repeat sequences required for transposition (Sherratt 1989). The *nptII* selectable marker gene in the mini-Tn1000 transposon is under the control of the *nos* (nopaline synthase) promoter and termination signals (Egener et al. 2002). In *Physcomitrella* the bacterial *nos* promoter allows strong *nptII* gene expression at levels comparable to the standard CaMV 35S promoter (Holtorf et al. 2002b).

The donor strain R2217 also harbours the plasmid pFDX3957, which contains the bacterial transposase gene *tnpA* and the kanamycin resistance gene *kan*. TnpA is inducible through an IPTG-responsive promoter. IPTG induction activates expression of the bacterial transposase gene leading to transposition of mini-Tn1000::nptII. Upon transposition a cointegrate structure is formed between the plasmids pUCMinIV and pFDX3951. Conjugation of the donor with the recipient bacterial strain R1037 leads to transfer of the cointegrate structure into the recipient cell. Upon transfer the cointegrate is resolved by activity of the IPTG-induced resolvase gene *tnpR* of the recipient cell. After conjugation and resolution cells are selected on an appropriate antibiotic-containing medium and plasmid DNA prepared from those cells harbouring the pUCMinIV::mutagenised cDNA. Resulting plasmid DNA batches are digested with the rare-cutting restriction enzyme *SdaI* yielding linear DNA that is subsequently used for transformation of *Physcomitrella* protoplasts.

This transposon shuttle mutagenesis method has the advantage that anonymous cDNA populations can be rapidly manipulated *en masse* without the need for sequence information. From a single moss cDNA library a multitude of chimeric

cDNA molecules, which harbour a selection cassette that disrupts the coding sequence, can be generated.

Shuttle mutagenesis leads to randomization of transposon-integration within the cDNA sequences, as there are no preferential sites for transposon insertion (Miller 1992). Non-biased transposon integration into moss cDNAs is illustrated in Figure 3. In this case 72 independent transposon insertion sites were mapped after mutagenesis by DNA sequence analysis of random individual clones from a single representative moss cDNA clone S_PP015059353 (808 bp). The experiment demonstrates that there is no apparent bias not only for the insertion of the transposon into a certain region of the cDNA but also for orientation of the transposon relative to the cDNA. In addition, the transposition of the mini-Tn1000 transposon into the cDNA occurred only once per plasmid molecule. Very few insertions into the pUCMinIV vector backbone were discovered.

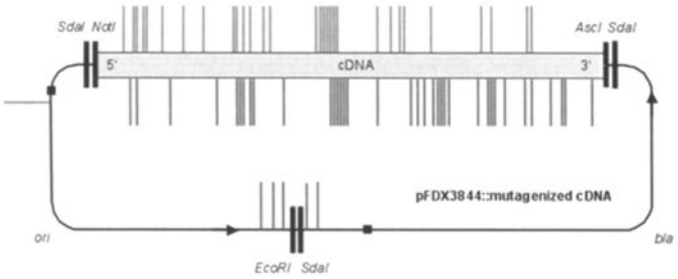

Figure 3
Distribution of transposon insertion sites in a representative *Physcomitrella patens* cDNA. Scheme depicting the orientation of 72 independent transposon insertion sites in a representative moss cDNA clone (ID: S_PP015059353; 808 bp). Prior to mutagenesis the cDNA was cloned into the minimal vector pUCMinIV. This defined plasmid was subjected to shuttle mutagenesis (see Fig. 2), and the transposon insertion sites for 72 resulting clones were mapped by DNA sequencing to assess the distribution of insertions. 41 insertions were found to be in "forward" orientation (*nptII* resistance marker on transposon and *bla* marker on vector in same orientation). "Forward" orientations are indicated by verticle lines below the cDNA, 31 "reverse" insertions by verticle lines below the cDNA. Most of the insertions (66 / 72, corresponding to 92%) occurred throughout the cDNA, without apparent strong preference for a certain insertion site or orientation. The *ori* and *bla* gene and restriction sites for *EcoRI*, *SdaI*, *AscI* and *NotI* are shown.

4.1 Fast forward production of single KO constructs

The usefulness of targeted gene knockouts for the elucidation of *Physcomitrella* gene function has recently been demonstrated (Strepp et al. 1998, Girke et al. 1998, Girod et al. 1999, Imaizumi et al. 2002, Koprivova et al. 2002, Zank et al., 2002). In these cases known genes were investigated for their function. For our program, based on the developed high-throughput transformation platform, the initial aim was to produce knockout plants for 80 candidate genes. Suitable target genes were identified on the basis of sequence homology to know3n genes by our industrial collaborator BASF Plant Science. These genes were chosen for their involvement in a variety of biochemical, developmental and signalling pathways. For each gene the corresponding EST clone was retrieved from the clone repository and disruption cassettes constructed manually. The resultant disruption constructs harboured the *nos-nptII-nos* selection

cassette in an appropriate position in the central part of the cDNA. The routine transformation protocols (see subchapter 2.5.) allowed for the generation of close to 5,000 gene mutants in total, 4,898 transformants of which were analysed for knockout insertions by a high-throughput PCR-protocol using primers that flanked the predicted genomic insertion site (Fig.4). An undisrupted wild type gene yields a short, abundant PCR product whereas a much longer PCR fragment can be anticipated in the knockout plants. This longer fragment is often amplified inefficiently and thus a loss of the wild-type PCR band was taken to indicate a disrupted gene. Plants identified using this method of gene disruption were confirmation of the insertion event by Southern blot analysis. Preliminary results involving 61 candidate genes indicate that the average gene targeting (GT) efficiency was 27% . The highest GT efficiency observed for a single construct was 88%, however, a GT efficiency of less 5% was obtained for almost one fifth (13 of 61) of the target genes (T. Egener and J.M. Lucht, personal communication). The low GT frequency in 20% of the tested knockout constructs may be explained by the possibility that gene disruption affects viability in these cases or an unfavourable exon-intron structure may be formed within the genomic locus. As the genomic locus for the majority of the 61 candidate genes is not known, we cannot rule out nor prove either of these possibilities. We have observed, however, that gene targeting efficiency increases when using long continuous stretches of genomic sequence for the construction of the knockout transgenes (J. Granado personal communication, H. Holtorf unpublished results). Interestingly, there was no strong correlation between GT efficiencies and length of the cDNA stretches used in the

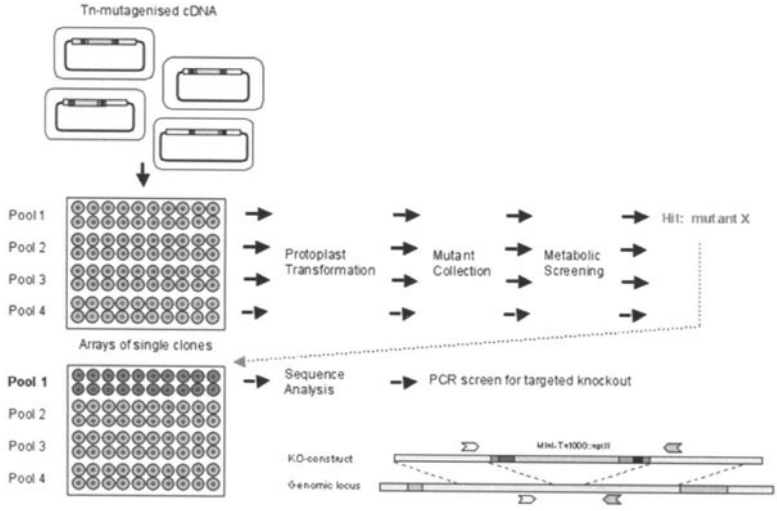

Figure 4 Production and analysis of mutants produced with batches of defined Tn-mutagenised cDNAs. Scheme outlining the production of moss mutants from batches of defined *Physcomitrella* cDNA clones. Pools containing 20 randomly selected cDNA clones are used for preparation of Tn-mutagenised plasmid DNA for protoplast transformation. After phenotypic screening mutants which show deviating metabolic profiles are selected for molecular analysis. Plasmid DNAs from clones of the corresponding pool are sequenced. Appropriate oligonucleotide primers are selected to flank the mini-Tn1000 transposon. Mutant DNA is screened for knockout insertions.

knockout constructs. The study demonstrates that routine, high-efficiency gene targeting using cDNA constructs is possible in *Physcomitrella*. A large number of candidate gene knockouts can be achieved in a relatively short period of time and mutants can be efficiently screened by PCR assays based on EST sequence fragments.

4.2. High throughput transformation protocol

The large-scale gene-function correlation study we have established is aimed at producing a collection of transgenics with insertion mutations that saturates the population of expressed *Physcomitrella* genes. To achieve this a robust transformation protocol that ensures reproducible output efficiencies and a high rate of knockout mutants must be developed. In this case, mutagenised cDNA for transformation was produced by transposon mutagenesis (Fig. 2) and the derived clones provided the transgene DNA. Plasmid DNA was prepared and transformed in batches. The mutagenised cDNA was integrated into the genome, preferentially by homologous recombination. The use of cDNA batches derived from the three different tissues mentioned earlier makes it reasonable to expect that all expressed moss genes will be disrupted with a targeting construct.

The polyethylene glycol (PEG)-based transformation protocol for direct DNA transfer to protoplasts (Strepp et al. 1998) had to be adapted in order to efficiently produce the required 80,000 mutants. Standard transformations utilize $3x10^5$ cells and 20 µg of linearized plasmid DNA. Transgenic plants are regenerated and selected on medium supplemented with various metabolites to facilitate survival of metabolic mutants (see subchapter 2.7.). The new protocol generated stable transformants in six weeks, using an initial selection on 25 µg/ml G418 for 2 weeks, a non-selective release step of 2 weeks, and an additional 2 weeks of G418 selection. Two rounds of selection on the antibiotic G418 are necessary to eliminate unstable transformants (Ashton et al. 2000).

Large-scale production of protoplast material was achieved by using a bioreactor culture protocol (see subchapter 2.1.) and improved culture conditions for optimal regeneration of protoplasts (Schween et al. 2002a). The transformation protocol was further optimized with respect to DNA conformation and pre-culture conditions which improved rates for plant output and gene targeting efficiency (Hohe et al. 2002). We compared transformation efficiencies for both circular and linear plasmids and demonstrated a difference in the yield of stable transformants. Transformation using circular DNA molecules resulted in only 0.2 % surviving plants after the second round of G418-selection, contrasting with 16 % for linear DNA. The low efficiency of stable transformation with circular DNA maybe a reflection of the finding that plasmids can be kept as extra-chromosomal replicating elements which rarely integrate into the genome (Ashton et al. 2000).

Disruption constructs from three different genes were used to generate optimal conditions for the pre-culture of plant material for transformation. Pre-culture conditions in the bioreactors were varied by supplementation with either ammonium tartrate or by controlling the pH (set point 4.5) of the culture solution. Both protocols resulted high protoplast yields but the transformation efficiencies were eight times higher in protoplasts derived from the pH controlled cultures. A pH-controlled bioreactor pre-culture enabled the production of 48 stable transgenic moss plants per

transformation. GT efficiencies mutants were measured by PCR-based screening procedures.

With these procedures it is possible to perform 18 independent transformations per day with a capacity of 9 plants per transformation, i.e. 160 plants per day. The output is highly reproducible and does not vary significantly for any of the mutagenised cDNA libraries that were used for transformation (A. Hohe personal communication). This high throughput transformation protocol will ensure the production a mutant collection of 80,000 plants, saturating the expressed genome of *Physcomitrella.*

Initially the overall number, identity, and nucleotide sequence of the different Tn-mutagenised cDNAs that made up a single batch of constructs used for transformation was not determined. This strategy hindered later characterization of the resulting transgenic plants. To correct this each batch of DNA was confined to 20 arbitrarily selected cDNA clones from a single cDNA library. The 20 clones constitute a single DNA pool. Each DNA pool and the individual clones contained therein are numbered and stored in a repository (Fig. 4). cDNA clone identification by sequence analysis is performed only when a mutant with an interesting phenotype emerges (see subchapter 2.8.) and the sequence of the 20 transgenes can be used to design oligonucleotide primers for PCR screening. Such cataloguing and tracking of clones enables a more rapid analysis of the gene lesion that causes the mutant phenotype.

4.3 Mutant quality control

To produce a mutant collection of 80,000 plants is not only challenging with respect to the number of plants that have to be produced but also with regards to mutant quality. Even routine protocols that underlie the production of standardized material cannot guarantee the absence of genetic changes that may arise during cell culture.

Two important quality traits of the generated plants are of special interest. First, all plants that have been selected on G418-containing medium must be tested to confirm the presence of the *nptII* selectable marker gene. Only these mutants can be considered as stable transformants. Second, the nuclear DNA content of all mutants must be determined in order to detect those plants that may arise from protoplast fusion events which double the chromosomal DNA content. Stable integration of the *nptII* selectable marker gene can be determined by a one-step genomic DNA extraction and a subsequent PCR assay (Schween et al. 2002c). Some of the mutants exhibit slow growth during the regeneration process which necessitated an adaptation of the DNA isolation protocol to accommodate small amounts of plant material (1 to 5mg). Furthermore, the genomic DNA preparation had to be fast and simple enough to to analyse 800 to1,000 plants per week (Schween et al. 2002c). The PCR protocol was optimized by addition of the linear multivalent polyamine spermidine and polyvinylprrolidone (PVP) to negate the inhibitory effects of polyphenols. The described PCR screen is fast and accurate and can be used to analyse 300-400 *Physcomitrella* plants per day. In an analysis of 15,000 putative *Physcomitrella* transformants, 13.7% did not yield an *nptII*-specific PCR product. Such plants were subjected to a third round of selection on G418-containing medium. The majority of

these plants survived the additional antibiotic treatment and were considered to be stable mutants. Multiple rounds of alternating selection and release treatments will only allow stably transformed plants to survive (Ashton et al. 2000). As a result 98% of plants were determined to be stable transgenics.

The PEG-transformation protocol occasionally yields polyploid plants, which may arise from protoplast fusion during the transformation process. To ensure a constant output of haploid mutants it is necessary to determine the nuclear DNA content of each transformant. Each mutant plant is required to be subjected to flow cytometric analysis (FCM). For routine mutant analysis by FCM the haploid genome of *Physcomitrella* was determined to be 511 Mbp in size (Schween et al. 2002c). In principle FCM analysis can be achieved with either protoplasts or isolated nuclei (Galbraith 1989; Ulrich and Ulrich 1991). For high throughput purposes a protocol described by Ulrich and Ulrich (1991) was adapted for *Physcomitrella*. This procedure required only 30 mg of fresh mutant tissue, from which intact nucleisuspensions were prepared by maceration of the material with a razor blade in a glass petri dish containing 2.5 mL of a DAPI-containing buffer. FCM measurements revealed that, using the standardized transformation and regeneration process, only 5% of 16,000 mutants tested were polyploid (Egener et al. 2002).

4.4 Screening for mutant phenotypes

Following the classical principle of forward genetics, all mutants produced from a Tn-mutagenised cDNA library were subjected to three different phenotypic screens: physiological, morphological, and metabolic. Screens for nutritional mutants of *Physcomitrella* have been described (Ashton and Cove 1997) and thus served as a model for an initial search for physiological alterations induced in moss mutants. To accomplish this mutant plants were regenerated and selected on both minimal and supplemented media. The supplemented media was Knop medium supplemented with MS microelements (Murashige and Skoog 1962), 4 mg/l myo-inositol, 2.8 mg/l choline chloride, 1 mg/l nicotinic acid, 0.5 mg/l thiamine-HCl, 0.25 mg/l pyridoxine, 0.01 mg/l biotin, 0.25 mg/l p-aminobenzoic acid, 1.9 mg/l Ca-D-pantothenate, 0.015 mg/l riboflavine, 6.76 mg/l adenine, 3.84 mg/l Na-palmitinic acid, 250 mg/l peptone, 920 mg/l ammonium tartrate and 50 g/l glucose.

Only 7.2% (1,163 of 16,203) of the regenerated plants experienced growth retardation when cultured on minimal medium in contrast to normal growth rates on the supplemented medium. Wild-type plants grew well on both media. The slow-growth phenotype on minimal medium suggests a metabolic defect that is rescued by supplying metabolites to the medium. A retarded growth during the culture on both minimal and supplemented meda, observed in 3.1% (508 of 16,203) of the transformants, suggests either the absence of a required metabolite or an unknown defect that is reflected in a low growth phenotype.

To screen for developmental phenotypes, mutants were scored, relative to the wild-type, for visible morphological aberrations and growth behaviour. Digital images and microscopic observations for each mutant were recorded and archived. Mutants were scored directly following regeneration and during selection on supplemented Knop medium for 11 weeks followed by growth on minimal Knop medium for 8

weeks. Under these conditions, the filamentous protonema underwent a developmental switch resulting in the formation of a three-faced apical meristem bud. This bud differentiated into the gametophores that consist of a stem and associated leafy structures. Eventually the gametophores carried sex organs. Deviations from this normal developmental pathway were observed in 16.2% (2,631 of 16,203) of the gene-disruption library transformants. These mutants exhibited one or more of the following features: changed structure and colour of the moss plant, unusual coverage of the plant by gametophores, disturbed shape and uniformity of leaves, and altered arrangement of cells within the leaves (Egener et al. 2002). Control plants were continuously scored to ensure that the suspect phenotypes originated from the insertion of transgenes into the nuclear genome and not from some unknown factor experienced in culture. Morphological deviations did not occur spontaneously within 350 wild-type plants and in less than 1% of the 400 plants derived from protoplasts that had undergone a mock-transformation without DNA and regenerated in the absence of the antibiotic G418. In addition to scoring growth requirements and morphology the ploidy level was determined for each mutant (see subchapter 2.6.). There did not appear to be any strict correlation between any of these three parameters but various combinations of characteristics were observed (Egener et al. 2002).

Egener et al. (2002) describe an impressive collection of physiological and morphological mutant phenotypes uncovered by the described screening protocol. The frequency of altered phenotypes in the mutant collection was considerably higher than the fewer than 2% reported for Arabidopsis confirmed knockout-mutants (Bouche and Bouchez 2001), which might be attributable to haploidy and a lower degree of gene redundancy in *Physcomitrella* (Rensing et al. 2002). The reported phenotypes leads to the conclusion that a wide spectrum of morphological and developmental alterations do occur in *Physcomitrella* plants transformed with the Tn-mutagenised gene-disruption library.

Visible phenotypes that are revealed during physiological and morphological screens are scored with relative ease, however the loss of a functional gene may result in subtler and less visible alterations in the biochemical composition of a mutant. Certain metabolites can be lost; others in turn may increase in amount or be modified to give rise to new compounds. To identify metabolic mutants in the collection a third screen will focus on the comparative analysis of the composition of metabolites within the *Physcomitrella* mutant collection. This screen will take advantage of new technologies, developed by Metanomics, Berlin (www.metanomics.de), that have emerged in the field of metabolomics. Such high-throughput approaches aim to identify gene function on the basis of a non-biased, simultaneous and rapid analysis of the metabolome (Trethewey et al. 1999; Trethewey 2001). The metabolic status of each mutant will be profiled with respect to a wide range of chemical constituents. Changed levels of certain compounds can reflect the presence or absence of gene products involved in synthesis or break-down of the compounds. Such analyses are predicted to aid in the identification of the immediate biochemical function of plant genes (Roessner et al. 2002).

Convincing examples of a link between a *Physcomitrella* gene knockout and a clear biochemical phenotype have been demonstrated (Girke et al. 1998, Zank et al. 2002, Sperling et al. 2002). These early studies demonstrated that *Physcomitrella* is a source for novel elongase and desaturase genes in the plant fatty acid biosynthetic

pathway from significant changes of fatty acid profiles in knockout mutants. The alteration in fatty acid profiles allowed the clear assignment of a function to the genes under investigation.

Novel insights into another plant biochemical pathway were gained by "knocking out" the *Physcomitrella apr* gene. The phenotype of mutant the revived interest in an old proposed route for sulfate assimilation in plants (Koprivova et al. 2002).The expectation of the metabolic analysis of the *Physcomitrella* mutant collection is that, as for the three cases mentioned above, a link between various biochemical phenotypes and the corresponding gene lesions will be established.

4.5. Molecular analysis of mutants

The overall purpose of the high-throughput, large-scale *Physcomitrella* genome project is to identify a large number of tagged *Physcomitrella* mutants that exhibit a commercially interesting phenotype. The high rate of homologous recombination, in combination with the *nptII* marker gene-tag, will enable rapid access to the disrupted gene sequence. The tagged loci are recoverble using a set of standard techniques that were adapted for molecular analysis of *Physcomitrella*. Efficiently screens for knockout insertions utilizing high-throughput PCR methods were developed (Schween et al. 2002b). In addition, by combining a newly developed DNA extraction protocol (Schlink and Reski 2002) with a non-radioactive detection method (Egener et al. 2002) the ease, speed and accuracy of gene knockout analysis by Southern blotting of genomic DNA has been improved.

Egener et al. 2002 assessed the overall quality of mutants, with respect to the number of integrated *nptII*-marker cassettes and integration loci. Molecular analyses were initially performed using Southern blots of isolated genomic DNA from randomly chosen transformants probed with suitable target sequences. This was achieved by digestion of the genomic DNA with *Pvu*II, to obtain an estimate for the number of gene-disruption constructs integrated in the transformed moss plants, since the *nptII* gene has a single *Pvu*II recognition site within its coding sequence. Restriction of this sequence results in the formation of two hybridisation bands, which are detected by a *nptII*-derived radioactive probe. The derived hybridisation bands differ in length from mutant to mutant depending on the nature of the adjacent sequences. These sequences can be derived either from genomic sequences or from transgenic cDNA fragments. Such a Southern blot analysis of 154 transformants revealed that the average number of *nptII* genes per mutant was approximately 10 (K. Schlink, personal communication).

To evaluate the average number of insertion sites per mutant genomic DNA of 16 transformants was digested with several restriction enzymes that cut genomic *Physcomitrella* DNA frequently but do not cut within the *nptII* cassette, and only rarely within the cDNA sequences carried by the transforming DNA (Egener et al. 2002). Analysing the *nptII*-probed restriction patterns discriminates between independent integration events at different genomic positions and co-integrations at the same locus. The majority of fragments, detected after hybridisation, were more than 20 kb in size; only a few distinct *nptII*-hybridising bands were observed. This result suggested that most of the Tn-mutagenised transgenic cDNA molecules are co-integrated in close vicinity to each other at up to 3 independent genomic loci. This is in concurrence with

observations that *Physcomitrella* integrates multiple transgene copies at one genomic locus, i.e. integration loci originating from a single homologous or illegitimate recombination event can carry multiple concatameric copies of transgenes (Schaefer and Zryd 1997, Nishiyama et al. 2000). Indeed, PCR amplification of several transgene sequences between the *nptII*-cassettes and subsequent sequence analysis of the products confirmed tandem integrations of different gene disruption constructs in close vicinity (Tanja Egener, Hauke Holtorf, unpublished results).

For gene-function assignment, however, the exact structure of the targeted locus is not of primary interest. Only a knowledge of the disrupted gene sequence at the border of the integration locus is relevant and needs to be determined. The challenging task is to employ suitable and rapid methods for identification of sequences that flank the insertion site. Various PCR-based methods can be used for this purpose (Ochman et al. 1993). In this project the borders were determined using a combination of PCR-based and plasmid rescue methods. For example, inverse PCR (iPCR, Hartl and Ochman 1994, Triglia et al. 2000), a combination of circularization of genomic fragments after restriction digestion and inversely oriented amplification, is one procedure used in the determination of genomic sequences that flank an inserted marker gene. Inverse PCR has been successfully used with moss genomic DNA and efficiently amplified sequences flanking the *nptII* gene. However, fragments generated by this method were small, making it difficult to unequivocally assign fragments to the gene insertion border.

A more suitable method is Thermal Asymmetric InterLaced PCR (TAIL-PCR) as described by Liu et al. 1995. TAIL-PCR is based on the combination of specific nested primers of (for example) a marker gene, with a set of nested, degenerate primers which can anneal in the genomic regions bordering the integrated transgene. TAIL-PCR is a very sensitive method that can be applied to large and complex genomes (Liu et al. 1995). This procedure has been adapted to isolate and sequence flanking genomic sequences from *Physcomitrella* DNA.

The presence of concatameric structures at the insertion locus, however, necessitates thorough inspection of sequence data produced from amplified fragments. This problem arises from the fact that sequences from inside the concatamer, that neighbour a *nptII*-cassette, are likely to be amplified. This is also true for Arabidopsis, where TAIL-PCR has been reported to be a versatile tool for isolation and mapping of T-DNA insert junctions and other border sequences (Liu et al. 1995, He and Gan 2001, Okamoto and Hirochika 2001). TAIL-PCR will be successful as long as *nptII* marker gene sequences lie in the vicinity of the disrupted genomic locus. TAIL-PCR specific products can be used for direct sequencing (Mazars and Theillet 1996) and can be immediately used to test the integrity of genomic loci by Southern blotting.

Plasmid rescue is also a powerful tool for cloning plant sequences that flank a selection marker gene (Mathur et al. 1998, Nakazawa et al. 2001). A suitable plasmid vector has been developed for the *Physcomitrella project* that enables the trapping of sequences that flank the mini::Tn1000 transposon. The trap vector contains an incomplete N-terminal fragment of the *nptII* marker gene under the control of a bacterial promoter. To clone border sequences, genomic DNA of a mutant plant is digested with appropriate restriction enzymes and ligated into the trap vector. Upon selection with kanamycin, only a restored marker gene will confer antibiotic resistance to the bacteria. The transgenic nucleotide sequence flanking the *nptII* gene can then be

determined by standard sequence analysis. Using this method several border sequences have been cloned (Tanja Egener, personal communication).

Plasmid rescue and TAIL-PCR protocols have successfully been used to assign border sequences to insertion sites in tagged mutants of Arabidopsis (McElver et al. 2001, Budziszewski et al. 2001). More than 25 genomic loci from *Physcomitrella* knockout mutants generated with known and anonymous disruption constructs have been isolated to date. Border sequence analysis was performed by comparing sequences of PCR products or fragments obtained by plasmid rescue with the public databases and our proprietary EST-database using BLAST (Altschul et al. 1997). Currently TAIL-PCR and plasmid rescue are the methods of choice to allow rapid isolation of tagged genomic loci in *Physcomitrella.*

Analysis of mutants from transformations, generated from defined cDNA pools (see subchapter 2.5.), is performed by the use the fast and comprehensive PCR-based screen to identify genes with a putative knockout insertion described here. Prior to the application of the protocols for isolation of tagged genes, each clone within the arrayed pool of clones is subjected to sequence analysis to identify the gene sequence and location of the transposon. It is the ability to identify the precise location of the transposon that is the advantage to using defined Tn-mutagenised cDNA batches: TO generate each mutant a defined pool of knockout constructs are used for transformation, and for each construct the position of the transposon is known, and only a maximum of 20 transgenes potentially integrate into the genome. The knowledge of the number and sequence of inserted transgenes allows for a much faster molecular analysis. Much like the candidate gene knockout approach (see subchapter 2.4.), primers are chosen such that they flank the transposon sequence of the disruption construct and genomic mutant DNA is subjected to PCR screening to identify the loss of an amplification product that is specific for the wild-type gene (Fig.4). All 20 transgenes can be rapidly screened for a knockout integration in the corresponding gene. The knowledge of the construct used to generate a disrupted wild-type gene can subsequently be used in the design of a Southern analysis. A band shift in the mutant relative to the wild-type is indicative of a knockout insertion at the gene in question. For gene isolation, genomic knockout mutant DNA can be subsequently analysed using TAIL-PCR and plasmid methods.

5. OUTLOOK

The moss *Physcomitrella patens* has emerged as an attractive model system for plant biology and functional genomic analysis. It shares many biological features with higher plants but has the unique ability for efficient homologous recombination within its nuclear DNA. In comparison to other plant models this advantage allows for precise genetic manipulation that can routinely produce targeted knockouts to study plant gene function. We have demonstrated that *Physcomitrella* is amenable to the large-scale functional genomics approache that allows the innovative creation, and efficient characterization, of plant genetic diversity to determine novel gene functions. The knowledge generated by such a model has the potential for application in crop plant improvement, especially in light of the high degree of genetic and physiological conservation between mosses and higher plants.

6. REFERENCES

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.,* 25, 3389-3402

Ashton, N.W., & Cove, D.J. (1977) The isolation and preliminary characterisation of auxotrophic and analogue resistant mutants of the moss *Physcomitrella patens*. *Mol. and Gen. Genet.,* 154, 87-95

Ashton, N.W., Champagne, C.E.M., Weiler, T., & Verkoczy, L.K. (2000) The bryophyte *Physcomitrella patens* replicates extrachromosomal transgenic elements. *New Phytol.,* 146, 391-402

Avila, P, & de la Cruz, F. (1988) Physical and genetic map of the IncW plasmid R388. *Plasmid,* 20, 155-157

Bouche, N. & Bouchez, D. (2001) Arabidopsis gene knockout: phenotypes wanted. *Curr. Opin. Plant. Biol.,* 4, 111-117

Boyd, P.J., Hall, J., & Cove, D.J., (1988) An airlift fermenter for the culture of the moss *Physcomitrella* patens. In *Methods in bryology*, Proc. Bryol. Meth. Workshop (Mainz, J. & M. Glime eds). Nichinan: *Journal of the Hattori Botanical Laboratory*, 41-45

Budziszewski, G.J., Lewis, S.P., Wegrich-Glover, L., Reineke, J., Jones, G., Schlater-Ziemnik, L., Lonowski, J., Nyfeler, B., Aux, G., Zhou, Q., McElver, J., Patton, D.A., Martienssen, R., Grossniklaus, U., Ma, H., Law, M., & Levin, J.Z. (2001) Arabidopsis Genes Essential for Seedling Viability: Isolation of Insertional Mutants and Molecular Cloning. *Genetics* 159, 1765-1778

Cove, D.J., Knight, C.D., & Lamparter, T.L. (1997) Mosses as model systems. *Trends Plant Sci.,* 2, 99-105

Egener, T., Granado, J., Guitton, M.C., Hohe, A., Holtorf, H., Lucht, J.M., Rensing, S., Schlink, K., Schulte, J., Schween, G., Zimmermann, S., Duwenig, E., Rak, B., & Reski, R. (2002) High frequency of phenotypic deviations in *Physcomitrella patens* plants transformed with a gene-disruption library. *BioMed. Cen. Plant Biol.*, 2, 6

Galbraith, D.W. (1989) Analysis of higher plants by flow cytometry and cell sorting. *Int. Rev. Cytol.* 116, 165-228

Girke, T., Schmidt, H., Zaehringer, U., Reski, R., & Heinz, E. (1998) Identification of a novel delta 6-acyl group desaturase by targeted gene disruption in *Physcomitrella patens*. *Plant J.* 15, 39-48

Girod, P.A., Fu, H., Zryd, J-P., & Viestra, R.D.(1999). Multiubiquitin chain binding subunit MCBI (RPN10) of the 26S proteasome is essential for developmental progression in *Physcomitrella patens*. *Plant Cell* 11, 1457-1472

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley. D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.L., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A., & Briggs, S. (2002) A draft sequence of the rice genome (L. ssp. *japonica*). *Science* 296, 92-100

Grimsley, N.H., Ashton, N.W., & Cove, D.J. (1977). The production of somatic hybrids by protoplast fusion in the moss *Physcomitrella patens*. *Mol. Gen. Genet.* 154:97-100

Hartl, D.L., & Ochman, H. (1994) Inverse polymerase chain reaction. *Methods Mol. Biol.*, 31, 187-96

He, Y., & Gan, S. (2001) Identical promoter elements are involved in regulation of the OPR1 gene by senescence and jasmonic acid in Arabidopsis. *Plant Mol. Biol.*, 47, 595-605

Hohe, A., & Reski, R. (2002) Optimisation of a bioreactor culture of the moss *Physcomitrella patens* for mass production of protoplasts. *Plant Sci.,* 163, 69-74

Hohe, A., Schween, G., & Reski, R. (2001) Establishment of a semicontinuous bioreactor culture of *Physcomitrella patens* for mass production of protoplasts. *Acta Horticulturae.* 560, 425-428

Hohe, A., Egener, T., Lucht, J.M., Holtorf, H., Reinhard, C., Schween, G,, & Reski, R. (2002) DNA conformation and plant pre-culture influence both, transformation efficiency and gene targeting efficiency in *Physcomitrella. Submitted.*

Holtorf, H., Guitton, M.C., & Reski, R. (2002a) Plant Functional Genomics. *Naturwissenschaften* 89, 235-249

Holtorf, H., Hohe, A., Wang, H-L., Jugold, M., Rausch, T., Duwenig, E., & Reski, R. (2002b) Promoter subfragments of the sugar beet V-type H^+-ATPase subunit c isoform drive the expression of transgenes in the moss *Physcomitrella. Plant Cell Rep.,* In press.

Imaizumi, T., Kadota, A., Hasebe, M., & Wada, M. (2002) Cryptochrome light signals control development to suppress auxin sensitivity in the moss *Physcomitrella patens*. *Plant Cell* 14, 373-386

Jenkins, G.I., & Cove, D.J. (1983) Light requirements for regeneration of protoplasts of the moss *Physcomitrella patens*. *Planta* 157, 39-45

Ko, M.S. (1990) An 'equalized cDNA library' by the reassociation of short double- stranded cDNAs. Nucl. Acids Res., 18, 5705-5711

Koprivova, A., Meyer, A.J., Schween, G., Herschbach, C., Reski, R., & Kopriva, S. (2002) Functional knockout of the adenosine 5'-phosphosulfate reductase gene in *Physcomitrella patens* revives an old route of sulfate assimilation. *J Biol. Chem.* 277, 32195-32201

Lal M (1984) The culture of bryophytes including apogamy, apospory, parthenogenesis and protoplasts. In: *The Experimental Biology of Bryophytes*, (Dyer A.F. &. Duckett J.G eds.). London, Academic Press: 97-115

Liu, Y.G., Mitsukawa, N., Oosumi, T., & Whittier, R.F. (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* 8, 457-463

Machuka, J., Bashiardes, S., Ruben, E., Spooner, K., Cuming, A., Knight, C., & Cove, D. (1999). Sequence analysis of expressed sequence tags from an ABA-treated cDNA library identifies stress response genes in the moss *Physcomitrella patens*. *Plant Cell Phys.*, 40,378-387

Mathur, J., Szabados, L., Schaefer, S., Grunenberg, B., Lossow, A., Jonas-Straube, E., Schell, J., Koncz, C., & Koncz-Kalman, Z. (1998) Gene identification with sequenced T-DNA tags generated by transformation of Arabidopsis cell suspension. *Plant J.* 13, 707-16

Mazars, G.R., & Theillet, C. (1996) Direct sequencing by thermal asymmetric PCR. *Methods Mol. Biol.* 65, 35-40

McElver, J., Tzafrir, I., Aux, G., Rogers, R., Ashby, C., Smith, K., Thomas, C., Schetter, A., Zhou, Q., Cushman, M.A., Tossberg, J., Nickle, T., Levin, J.Z., Law, M., Meinke, D., & Patton, D. (2001) Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana. Genetics* 159, 1751-1763

Miller J.H. (1992) A short course in bacterial genetics. *Cold Spring Harbor, N. Y.*: Cold Spring Harbor Laboratory Press.

Murashige, T., & Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue culture. *Physiol. Plant.* 15:473–497

Nakazawa, M, Yabe, N., Ichikawa, T., Yamamoto, Y.Y., Yoshizumi, T., Hasunuma, K., & Matsui, M. (2001) DFL1, an auxin-responsive GH3 gene homologue, negatively regulates shoot cell elongation and lateral root formation, and positively regulates the light response of hypocotyl length. *Plant J.*, 25, 213-221

Nishiyama, T., Hiwatashi, Y., Sakakibara, I., Kato, M., & Hasebe, M. (2000) Tagged mutagensesis and gene-trap in the moss, *Physcomitrella patens* by shuttle mutagenesis. *DNA Res.* 7, 9-17

Ochman, H., Ayala, F.J., & Hartl, D.L. (1993) Use of polymerase chain reaction to amplify segments outside boundaries of known sequences. *Methods Enzymol.* 218, 309-321

Okamoto, H., & Hirochika, H. (2000) Efficient insertion mutagenesis of Arabidopsis by tissue culture-induced activation of the tobacco retrotransposon Tto1. *Plant J.*, 23, 291-304

Rak, B., Reski, R., Zimmermann, S., Guitton, M.C., Duwenig, E., & Freund, A. (2001) Method for the mutagenesis of nucleotide sequences from plants, algae and fungi. *WO 01/38509 A1*

Rensing, S.A., Rombauts, S., Hohe, A., Lang, D., Duwenig, E., Rouze, P., Van de Peer, Y., Reski, R. 2002a) The transcriptome of the moss *Physcomitrella patens*: comparative analysis reveals a rich source of new genes. Available online: http://www.plant-biotech.net/Rensing_et_al_transcriptome2002.pdf.

Rensing, S.A., Rombauts,S., Van de Peer, Y., & Reski, R. (2002b): Moss transcriptome and beyond. *Trends Plant Sci..*, in press.

Reski, R., Faust, M., Wang, X.H., Wehe, M., & Abel, W.O. (1994). Genome analysis of the moss *Physcomitrella patens* (Hedw.) BSG. *Mol. Gen. Genet.* 244, 352-359.

Reski, R. (1998a). Development, genetics and molecular biology of mosses. *Bot. Acta.* 111, 1-15

Reski, R. (1998b) *Physcomitrella* and Arabidopsis: the David and Goliath of reverse genetics. Trends *Plant Sci.,* 3:209-210

Reski, R. (1999). Molecular genetics of *Physcomitrella*. *Planta*, 208, 301-309

Reski, R. & Abel, W.O. (1985) Induction of budding on chloronemata and caulonemata of the moss, *Physcomitrella patens*, using isopentenyladenin. *Planta*, 165, 354-358

Reski, R. Reynolds, S., Wehe, M., Kleber-Janke, T., & Kruse, S. (1998) Moss (*Physcomitrella patens*) expressed sequence tags include several sequences which are novel for plants. *Bot. Acta* 111, 145-151

Reutter, K., Atzorn, R., Hadeler, B., Schmülling, T., & Reski, R. (1998). Expression of the bacterial ipt gene in *Physcomitrella* rescues mutations in *budding* and *plastid division*. *Planta*, 206, 196-203

Reutter, K., & Reski, R. (1996) Production of a heterologous protein in bioreactor cultures of fully differentiated moss plants. *Plant Tiss. Cult. Biotech.* 2, 142-147

Roessner, U., Willmitzer, L., & Fernie, A.R. (2002) Metabolic profiling and biochemical phenotyping of plant systems. *Plant Cell Rep.*, In press

Rother, S., Hadeler, B., Orsini, J.M., Abel, W.O., & Reski, R. (1994). Fate of a mutant machrochloroplast in somatic hybrids. *J. Plant Physio*l., 143:72-77

Schaefer, D.G. (2001) Gene targeting in *Physcomitrella patens*. *Curr. Opin. Plant Biol*., 4, 138-141

Schaefer, D.G. (2002). A new moss genetics: targeted mutagenesis in *Physcomitrella patens*. *Ann. Rev. Plant Biol*., 53, 477-501

Schaefer, D.G., & Zryd, J.P. (1997) Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J*., 11, 1195-1206

Schaefer, D.G., Zryd, J.P., Knight, C.D., & Cove, D.J. (1991) Stable transformation of the moss *Physcomitrella patens*. *Mol. Gen. Genet*., 226, 418-424.

Schlink, K., & Reski, R. (2002): Preparing high-quality DNA from moss (*Physcomitrella patens). Plant Mol. Biol. Rep.*, 20, In press

Schween, G., Hohe, A., Koprivova, A., & Reski, R. (2002a): Effects of nutrients, cell density and culture techniques on protoplast regeneration and early protonema development in *Physcomitrella patens*. *J. Plant Physiol*, In press.

Schween, G., Gorr, G., & Reski, R. (2002b) Unique tissue-specific G2/M-arrest in *Physcomitrella. Submitted*

Schween, G., Fleig, S., & Reski, R. (2002c). High-throughput-PCR screen of 15,000 transgenic *Physcomitrella* plants. *Plant Mol. Biol. Rep.* 20, 43-47

Sherratt, D.J. (1989) Tn3 and related transposable elements: site-specific recombination and transposition. In: *Mobile DNA* (Berg, D.E., & Howe, M.M. eds) Washington, D.C: American Society for Microbiology:163-184

Sperling, P., Egener, T., Lucht, J.M., Reski, R., Cirpus, P., & Heinz, E. (2002): Identification of a Δ5-fatty acid desaturase from *Physcomitrella patens*. *Advanced Researches on Plant Lipids*, In press.

Strepp, R., Scholz, S., Kruse, S., Speth, V., & Reski, R. (1998). Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. *Proc. Natl. Acad. Sci. USA*. 95, 4368-4373

Stumm, I., Meyer, Y., & Abel, W.O. (1975) Regeneration of the moss *Physcomitrella patens* (Hedw.) from isolated protoplasts. *Plants Sci. Lett*. 5, 113-118

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796-815

Theissen, G., Münster, T., & Henschel, K. (1999) Why don't mosses flower? *New Phytol.,* 150, 1-8

Trethewey, R.N., Krotzky, A.J., & Willmitzer, L. (1999) Metabolic profiling: a rosetta stone for genomics? *Curr. Opin. Plant Bio.l* 2,83-85

Trethewey, R.N. (2001) Gene discovery via metabolic profiling. *Curr Opin.Biotech*., 12, 135-138

Triglia, T. (2000) Inverse PCR (IPCR) for obtaining promoter sequence. *Methods Mol. Biol*. 130, 79-83

Ulrich, I., & Ulrich, W. (1991) High-resolution flow cytometry of nuclear DNA in higher plants. *Protoplasma,* 165, 212-215

Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., & Zhang, X.et al., (2002) A draft of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296, 79-92

Zank, T.K., Zahringer, U., Beckmann, C., Pohnert, G., Boland, W., Holtorf, H., Reski, R., Lerchl, J., & Heinz, E. (2002) Cloning and functional characterisation of an enzyme involved in the elongation of Δ6-polyunsaturated fatty acids from the moss *Physcomitrella patens*. *Plant J.* 31, 255-268

MOLECULAR BIOLOGY AND GENOMICS OF THE DESICCATION TOLERANT MOSS *TORTULA RURALIS*.

ANDREW J. WOOD AND MELVIN J. OLIVER

Department of Plant Biology, Southern Illinois University—Carbondale, Carbondale IL 62901-6509 (AJW); Plant Stress and Water Conservation Unit, USDA-ARS, 3801 4th Street, Lubbock. TX, 79415(MJO)

Abstract. The molecular, cellular and biochemical responses of plants to water-deficit stress are the central, and long-term, interest of our respective research programs. We utilize the desiccation-tolerant moss *Tortula ruralis* (Hedw.) Gaerten., Meyer & Scherb. as an experimental model for studying post-transcriptional gene control, molecular & biochemical responses to abiotic stress, cellular repair mechanisms in plants and as a source of novel tolerance-associated genes. In this chapter we will introduce the molecular and biochemical mechanisms of desiccation-tolerance in mosses, describe the molecular and genomics tools that have been developed for *T. ruralis* (such as EST databases, cDNA libraries, and microarrays), and discuss the expression analysis of several cDNA clones (i.e. the rehydrins Tr288 & Tr213, the aldehyde dehydrogenases *Aldh7B6* & *Aldh21A1*, and the early light-inducible proteins *Elipa* & *Elipb*) that are associated with desiccation-stress.

1. INTRODUCTION

The mosses (Bryophatea, Musci) are a diverse and widely distributed group of land plants. Mosses can be found in a variety of habitats, but they have been particularly successful in colonizing "stressful" environments such as deserts, tundra, and substrate poor rocky outcrops. A significant number of moss species can exploit such water-limited niches because they have evolved the ability to tolerate desiccation and to grow when free water is available. Desiccation-tolerance, (i.e. the ability of cells to revive from the air-dried state) (Bewley 1979, Alpert and Oliver 2002) is common amongst mosses, although it is not uniformly distributed between species (Oliver et al. 2000a, Proctor and Tuba 2002). The genus Pottiaceae is comprised of many species that inhabit xeric (or "harsh") environments. One such species *Tortula ruralis* (Hedw.) Gaerten., Meyer & Scherb. (*Syntrichia ruralis* (Hedw.) F.Weber & D.Mohr), is the focus of this chapter and exhibits the desiccation –tolerance phenotype.

A.J. Wood et al. (ed.),
New Frontiers in Bryology: Physiology, Molecular Biology and Functional Genomics, 71–89.

Over the past 30 years *T. ruralis* has become an important experimental model for the study of vegetative desiccation-tolerance, organelle repair and recovery, and post-transcriptional gene control in plants (Oliver and Wood 1997, Oliver et al. 2000b). *T. ruralis* has been studied extensively at the molecular, biochemical, physiological and structural levels (see reviews in Bewley 1995, Oliver and Bewley 1997, Oliver and Wood 1997, Wood et al. 2000). We are convinced that *T. ruralis* will provide novel insights into the mechanisms of abiotic stress-tolerance in eukaryotes and provide an important phylogenetic comparison within mosses, bryophytes and the plant kingdom. In this chapter we will discuss the molecular and biochemical mechanisms of desiccation-tolerance, and outline the molecular and genomics tools that have been developed for *T. ruralis*.

2. THE PHYLOGENETICS OF DESICCATION-TOLERANT PLANTS

Recent synthetic phylogenetic analyses (Oliver et al. 2000a) suggest that vegetative desiccation tolerance was primitively present in the bryophytes (the basal-most living clades of land plants), but was lost in the evolution of tracheophytes. We postulate that the initial evolution of vegetative desiccation-tolerance was a crucial step required for the colonization of the land by primitive plants from a fresh water origin (Oliver et al. 2000a). We postulate that tolerance came at a cost, since metabolic rates are low in tolerant plants as compared to plants that don't maintain costly mechanisms for tolerance. Thus, the loss of tolerance might have been favored, along with the internalization of water relationships, as the vascular plants became more complex and entered into more competitive ecological niches. However, at least one independent evolution (or re-evolution) of desiccation-tolerance occurred in Selaginella and in the ferns. Within the Angiosperms, at least eight independent cases of evolution (or re-evolution) of desiccation-tolerance occurred. This phylogenetic evidence when combined with what we can deduce of the mechanisms of vegetative desiccation-tolerance leads to a substantive hypothesis as to the nature and progression of the evolution of this trait (Oliver et al. 2000a).

We hypothesize that the primitive mechanism of tolerance exhibited by the first plants (the cardinal mechanism, see 3. below) probably involved a constitutive level of cellular protection coupled with an efficient and active repair process similar to what we have described for modern day desiccation-tolerant bryophytes, in particular *T. ruralis*. This is consistent with the three criteria which must be satisfied in order for a plant to survive severe water deficits (Bewley, 1979): (1) damage must be limited to a repairable level, (2) physiological integrity must be maintained in the dehydrated state and (3) repair mechanisms must exist which function upon rehydration. As plant species evolved to fill the various niches available to them on dry land vegetative desiccation-tolerance was lost in preference for the advantages afforded to them by increased growth rates, morphological complexity, and mechanisms that conserve water whilst maintaining efficient carbon fixation. Genes that had evolved for cellular protection and repair were, in all likelihood, recruited for different but related process such as the response to water

stress and more importantly the desiccation-tolerance of reproductive propagules. We are thus hypothesizing that the mechanism of desiccation-tolerance exhibited in seeds evolved from the most primitive form of vegetative desiccation-tolerance. Once established in seeds the developmentally induced cellular protection system became available for induction in vegetative tissues by environmental cues that are related to drying. Thus the more recent modified vegetative desiccation-tolerance mechanism evolved from that programmed into seed development as certain species spread into very arid environments. More recently still, in response to the rigors of being dry in high irradiance habitats certain desiccation-tolerant monocots evolved the strategy of poikilochlorophylly to survive and compete in marginal habitats.

3. THE NATURE OF DESICCATION-TOLERANCE IN T. RURALIS

Vegetative desiccation-tolerance although relatively rare is represented by species that belong to most major classes of plants (Oliver and Bewley 1997, Oliver et al. 2000, Porembski and Barthlott 2000). Plants that exhibit vegetative desiccation-tolerance fall into two main categories; cardinal (meaning of basic importance; primary in an evolutionary sense) desiccation tolerant plants that can withstand the total loss of free protoplasmic water at any rate and modified (meaning derived from a non-primary mechanism) desiccation-tolerant plants that can only survive desiccation if water loss is slow. All cardinal desiccation-tolerant plants studied to date are of the less complex groups of plants; algae, bryophytes or lichens. The internal water content of these plants rapidly equilibrates to the water potential of the environment as they possess little in the way of water retaining morphological or physiological characteristics. The majority of the modified desiccation-tolerant plants are members of the more complex plant groups from ferns to angiosperms. These plants achieve slow drying rates by an array of morphological and physiological mechanisms that retard water loss to the extent required to establish tolerance

Studies aimed at the elucidation of vegetative desiccation-tolerance mechanisms at the molecular and genetic level, have centered around three plant species that represent the two classes of desiccation-tolerant plants. The modified desiccation tolerant angiosperms, *Craterostigma plantagineum* a southern African dicot and *Sporobolus stapfianus* an African desert grass, and the cardinal desiccation-tolerant bryophyte *Tortula ruralis* a North American moss (Bewley and Oliver 1992, Ingram and Bartels 1996, Oliver et al. 1997, Oliver and Bewley 1997). The prevailing evidence from these systems supports the idea of two basic strategies for vegetative desiccation-tolerance (Oliver and Bewley 1997). The modified desiccation-tolerant angiosperms appear to rely almost totally on an inducible cellular protection based mechanism that requires a certain amount of time to become established in the leaf tissues. The cardinal desiccation-tolerant bryophyte, however, appears to have a mechanism that relies on a certain constitutive level of cellular protection but more on a damage repair and reconstitution mechanism induced upon rehydration.

Desiccation of gametophytic tissues of *T. ruralis* results in a rapid decline in protein synthesis, as in all desiccation-tolerant and intolerant mosses tested so far (Bewley and Oliver 1992, Oliver and Bewley 1997). This loss of protein synthetic capacity is manifested in a loss of polysomes resulting from the run-off of ribosomes from mRNAs, concomitant with their failure to reinitiate protein synthesis. The rapid loss of polysomes during drying (under "natural" drying rates) and the apparent sensitivity of the initiation step of protein synthesis to protoplasmic drying leads us to the conclusion that the induction of synthesis of "protective" proteins during drying is highly unlikely. This is borne out by the observation that no new mRNAs are recruited into the protein synthetic complex even during slow drying (Oliver 1991). The fact that the moss survives rapid desiccation (even when desiccation is achieved in a few minutes in a lyophilizer), also indicates that an inducible protection mechanism is not necessary for survival.

The proposed mechanisms of desiccation tolerance differ between mosses and angiosperms (see 3. above and Phillips et al. 2002, Alpert and Oliver 2002). The well-studied angiosperm *C. plantagineum* utilizes, at least in part, a drying induced elevation of abscisic acid to trigger the accumulation of gene products that mediate the establishment of a cellular protection system prior to desiccation (Ingram and Bartels 1996). The stress-inducible protection system is comprised of two major components: the ABA-mediated synthesis of stress-proteins, such as dehydrins, early light-inducible proteins and aldehyde dehydrogenases, and the accumulation of active sugars such as sucrose. In sharp contrast, *T. ruralis* utilizes a constitutive protection system that employs key aspects of the inducible system characterized within angiosperms in addition to an active rehydration induced recovery mechanism (Oliver et al. 2000a). In hydrated *T. ruralis* gametophytes, dehydrin protein is readily detectable by Western blot analysis (Bewley et al., 1993) and sucrose content is approximately 10% g^{-1} DW (e.g. 292 mmol g^{-1} DW) (Bewley et al. 1978). *In vitro*, this amount of sugar is sufficient to offer membrane protection during drying (Straus and Hauser 1986). Interestingly, the relative amount of both dehydrin protein and sucrose is unchanged in response to desiccation or rehydration, and this lack of an increase in soluble sugars during drying appears to be a common feature of desiccation-tolerant mosses . The level of protection that is afforded by these mechanisms is at least sufficient to prevent any measurable damage to the important energy producing photosystems within the chloroplasts of this species (Proctor and Tuba. 2002).

Although much appears to be precluded during drying of gametophytic tissue of fully desiccation-tolerant bryophytes, there does appear to be some capacity to prepare for a future recovery event if drying rates are slow (4-6h). Using cDNA clones corresponding to *T. ruralis* transcripts that are preferentially translated during rehydration (see below and Scott and Oliver, 1994), it was determined that several "recovery" or rehydrin transcripts accumulate during slow drying (Oliver and Wood, 1997). These transcripts do not accumulate during rapid desiccation nor is their accumulation ABA-induced, indeed ABA is undetectable in this moss (unpublished data). Recent analysis of transcript accumulation during a time of metabolic decline has

revealed that transcripts are being sequestered in the dried gametophytes in messenger ribonucleoprotein particles (mRNPs) (Wood and Oliver 1999). The implication from this work is that the sequestration of mRNAs required for recovery hastens the repair of desiccation/rehydration-induced damage and thus minimizes the time needed to restart growth upon rehydration. These findings may also explain, in the absence of an inducible dehydrin and sugar response, the ability of T. *ruralis* to "harden" during recurring desiccation events (Shonbeck and Bewley 1981).

That *Tortula* requires a mechanism to repair cellular damage is brought out in ultrastructural studies of both dried and rehydrating gametophytic cells. Freeze fracture studies of dried *Tortula* cells (both rapid and slow-dried) clearly demonstrate that cellular integrity is maintained during drying (Platt et al. 1994). However, upon rehydration gametophytic cells undergo substantive and universal disruption of cellular integrity including what appear to be breaches to all membrane systems (see Oliver and Bewley 1984a for review). Nevertheless, the cells do not die, as do cells of sensitive species, but return to normal within 12 to 24h. Membrane leakage ceases within minutes of rehydration indicating a rapid membrane repair process that becomes active when water returns to the cells. These studies add credence to the notion that rehydration induced cellular repair is a major component of the desiccation-tolerance mechanism in *T. ruralis*.

Early work (see Bewley 1979 for review) established the ability of *T. ruralis* and other mosses to rapidly recover synthetic metabolism when rehydrated. In addition, although the pattern of protein synthesis in the first two hours of rehydration of *T. ruralis* is distinctly different from that of hydrated controls, novel transcripts were not made in response to desiccation (Oliver 1991, Oliver and Bewley, 1984b). Hence it was suggested that *T. ruralis* responds to desiccation by an alteration in protein synthesis upon rehydration that is in large measure the result of a change in translational control(s). In a detailed study of the changes in protein synthesis initiated by rehydration in *T. ruralis*, Oliver (1991) demonstrated that the synthesis of 25 proteins is terminated, or substantially decreased, and the synthesis of 74 proteins is initiated, or substantially increased, during the first two hours of hydration. The change in synthesis of these two groups of proteins, the former termed *hydrins* and the latter *rehydrins*, is not co-ordinately controlled and it takes a certain amount of prior water loss to fully activate the protein-based portion of the recovery mechanisms.

Changes were observed in the accumulation patterns of nearly all transcripts studied (Scott and Oliver, 1994) but these did not result in a qualitative change in the transcript population during desiccation or rehydration. Three distinct classes of cDNA clones can be discerned for *T. ruralis* based upon the relative recruitment of their corresponding transcripts into the polysomal fraction of hydrated, and rapid-dried (RD) rehydrated gametophytes (Oliver 1991). These classes are the constitutive cDNAs, cDNAs representing hydrin transcripts and cDNAs representing rehydrin transcripts (Scott and Oliver 1994). Hydrins are transcripts that are lost from polysomal fractions during rehydration, while rehydrins are transcripts that are preferentially found in rehydrating polysomal fractions. This change in the pattern of protein synthesis occurs

without a qualitative change in the pool of mRNA available for translation (Oliver and Bewley 1984b, Scott and Oliver 1994) suggesting that the transcriptional responses to effect the synthesis of stress related proteins that are common in stressed plants (Ingram and Bartels 1996), is not operating in the response of the moss to desiccation. The inference from this study is that the alteration in gene expression associated with rehydration is mediated mainly by an alteration in translational controls, at the level of mRNA selection. Perhaps, as suggested by the aforementioned studies of Wood and Oliver (1999), this control is mediated by the formation and release of rehydrin of mRNPs during drying and upon rehydration. Thus it appears that a major component of the desiccation tolerance mechanism of *T. ruralis* is the activation of pre-existing repair mechanisms in contrast to either pre-established or activated protection systems.

3. DEVELOPMENT OF MOLECULAR AND GENETIC TOOLS

Understanding how plant cells tolerate water loss is a vital prerequisite for developing strategies that can impact agricultural and horticultural crop productivity and survival under conditions of decreasing water availability. Most crops have a limited capacity for drought-tolerance and thus the genetic information necessary for expanding their drought-tolerance may not be exploitable or indeed present. We are convinced that understanding how plant cells tolerate the severest of water deficits, viz. desiccation, will offer novel perspectives and new insights into water stress tolerance mechanisms that cannot be obtained from drought-sensitive species. Using a genomics-based strategy, our long-term goal is to identify and determine the role of all genes involved in desiccation-tolerance in plants—from signaling to the establishment of tolerance.

3.1 Expressed Sequence tags (ESTs)

The structural analysis of genomes by large scale single-pass sequencing of randomly selected cDNA clones was pioneered using human brain tissue (Adams et al. 1991) and has been a powerful technique for the discovery of new genes in

Table 1. EST Summary

Type of Library	*Number of ESTs*	*Direction of Sequence*
I. cDNA Expression		
a. Slow-dried Polysomal mRNP	157[1]	3'
b. Rapid-dried Rehydrated	10,000[2]	5' and 3'
II. Subtractive Suppression Hybridization		
a. Slow-dried vs. Rehydrated	384[2]	5' and 3'
b. Rapid-dried Rehydrated vs. Rehydrated	384[2]	5' and 3'

[1]Wood et al. 1999, [2]M J Oliver unpublished results

Physcomitrella (Reski et al. 1998, Machuka et al. 1999, Chapters 1 & 3 this volume) and *Tortula* (Wood et al. 1999). In order to study those genes sequestered as mRNPs in desiccated *T. ruralis* gametophytes, we constructed a cDNA expression library using polysomal RNA from desiccated gametophytes and generated 152 EST sequences (Wood et al. 1999, Table 1). Only 29% (44/152) were significantly similar to DNA and/or amino acid primary sequences from other organisms. These small-scale EST sequencing projects established that moss genes/cDNAs are similar to angiosperms in terms of codon usage and are a valuable source of novel plant genes. Because the 152 slow-dried polysomal ESTs were generated from sequencing of 3' ends, we were able to characterize the putative cis-acting elements associated with the formation of mRNA 3'-ends (Wood et al. 2000b). Using this database over the past 4 years, we have employed 5' and 3'

Table 2. Inventory of Tortula ruralis cDNAs

cDNA	*Accession Number*	*Putative Identity/Orthologue*	*EST*
Tr155[1]	D63917	Peroxiredoxin	na
Tr213[1]	U29162	Polyubiquitin	na
Tr280[1]	U40818	Prolyl endopeptidase	na
Tr288[1]	AF275946	Dehydrin-like	na
Aldh7B6[2]	AY034889	Aldehyde dehydrogenase	na
Aldh21A1[3]	AY034888	Aldehyde dehydrogenase	AI305018
Elipa[4]	AY034890	Early-light inducible	AI304975
Elipb[4]	AY034891	Early-light inducible	AI665857
Rps3A[5]	AF093109	Ribosomal protein	AI305094
Rpl15[6]	AF230646	Ribosomal protein	AI305075
Rps14[7]	AF108724	Ribosomal protein	AI305056
Rps16[7]	AF108725	Ribosomal protein	AI305072
Rpl23[7]	AF108726	Ribosomal protein	AI305082
TrDr1[8]	AI304977	*psbI*	AI304977
TrDr2[8]	AI305064	desiccation stress cDNA	AI305064
TrDr3[9]	AY168202	Novel membrane protein	AI304984
TrDr4[10]	AF157017	*PsaG*	AI305105
TrDr5[11]	AF093108	Histone H3	AI305084
TrWD40A[9]	AF470350	WD40 protein	AI305059
Vac1[12]	AF303372	V-ATPase c subunit	AI305045
TrRPT2[13]	AF432345	26S Proteosome subunit IV	AI305024

[1]Scott and Oliver 1994, [2]Chen et al. 2002b, [3]Chen et al. 2002c, [4]Zeng et al. 2002, [5]Duff et al. 1999b, [6]Zeng and Wood 2000, [7]Wood et al. 2000a, [8]Triwitayakorn and Wood 2002, [9]AJ Wood, unpublished results, [10]Wood and Duff 1999, [11]Duff et al. 1999, [12]Chen et al. 2002a, [13]Chen and Wood 2003.

RACE (Rapid Amplification of cDNA Ends) to characterize full-length EST-derived cDNAs for a wide array of genes (Table 2).

EST analysis has been expanded to incorporate Subtractive Suppression Hybridization (SSH) cDNAs as well as ESTs derived from a rapid-dried rehydrated gametophytic tissue cDNA library. SSH technology is a recently developed PCR-based method for enriching tissue- or phenotype-specific cDNAs (Diatchenko et al., 1996) that we have used in an attempt to concentrate our focus on rehydrin mRNAs, both stored within mRNPs in the dried state and preferentially recruited into the active protein synthetic complex upon rehydration. We have developed SSH cDNA libraries where single stranded cDNAs derived from transcripts isolated from the polysomal fractions of both slowly dried and rapidly dried rehydrated gametophytes were subtracted against an excess of single stranded cDNAs derived from hydrated gametophyte transcript populations. The SSH protocol incorporates two subtractive hybridization steps that, along with the suppression PCR component, generates populations of cDNA clones enriched for sequences specific to the target transcript populations. These cDNA clones are generally not full length, from 300 to 600bp in length, and are biased towards the 3'-end of the target transcripts. Although we have a collection of 5000 clones from each of the two SSH libraries we have sequenced only 384 clones from each as yet. Of these 768 SSH/ESTs only 9% were significantly similar to DNA or predicted primary sequences from other organisms (Oliver, unpublished data). This low degree of similarity with sequences deposited in the public databases is probably the result of the shortness of the clones and their bias towards the 3'end of the transcripts rather than the possible uniqueness of the *Tortula* rehydrin genes (although this is a possibility). In contrast to this, we have sequenced (both 5' and 3') 10,000 ESTs, randomly isolated cDNA clones, from a library representing transcripts present during the initial phases of recovery following rehydration of rapid-dried gametophytes. Of these sequences 49% show significant similarity to sequences present in the current Genbank databases. The sequence data for these 10,000 ESTs are publicly available through the NCBI Genbank website. These EST resources will serve as the basis of expression, phylogenetic and genome level analyses in *Tortula.*

3.2 DNA Microarrays

Initial expression profile studies using cDNA microarrays are underway utilizing both SSH/ESTs and ESTs derived from the rehydrated gametophyte library. We have developed a small microarray containing triplicates for each of the 768 sequenced SSH clones along with appropriate control cDNA and several of the full-length cDNA clones described in Table 2. We are in the process of establishing a unigene collection of clones derived from the 10,000 rehydration EST collection. Our initial analyses of both the 5' and 3' sequences indicate that the collection represents approximately 3,500 individual *Tortula* genes. This unigene set will form the basis of a "desiccation-tolerance microarray" for expression studies in *Tortula.*

4. EXPRESSION ANALYSIS OF DESICCATION STRESS-RELATED cDNAs

In a series of publications over the past 10 years, we have hypothesized that genes essential to recovery and cellular repair are preferentially expressed upon rehydration of desiccated gametophytes (Scott and Oliver 1994, Oliver and Bewley 1997, Oliver and Wood 1997, Wood and Oliver 1999, Duff et al. 1999, Oliver et al. 2000a, b). We have further postulated that constitutively expressed transcripts consist of both important housekeeping genes and those genes that maintain the constitutive desiccation-tolerance protection system(s). We have used RNA gel blot analysis to characterize a number of cDNAs at the transcriptional level including the rehydrin Tr288 Velten and Oliver 2000), the aldehyde dehydrogenases *Aldh7B6* (Chen et al. 2002b) & *Aldh21A1* (Chen et al. 2002c), and the early light-inducible protein cDNAs *Elipa* & *Elipb* (Zeng et al. 2002). The accumulation of steady-state mRNA transcripts has been analyzed in response to a desiccation/rehydration cycle in both total RNA & Polysomal RNA fractions. In our experiments, we utilize gametophytic tissue that is hydrated , desiccated (either slowly or rapidly) or rehydrated. Hydrated moss is obtained after a 24 h rehydration period (14°C, 50 μE $m^{-2}s^{-1}$). Desiccated moss is obtained by drying over or a stirred saturated solution of sodium nitrite (slow-dried, SD) or over activated silica gel (rapid-dried, RD) as described (Chen et al. 2002a, b). Rehydrated moss is obtained by the addition of ddH_2O to desiccated moss (14°C, 50 μE $m^{-2}s^{-1}$). In this section, we will describe the characterization and expression of several key genes classes that are postulated to be active in the cytoplasm and chloroplast (Figure 1).

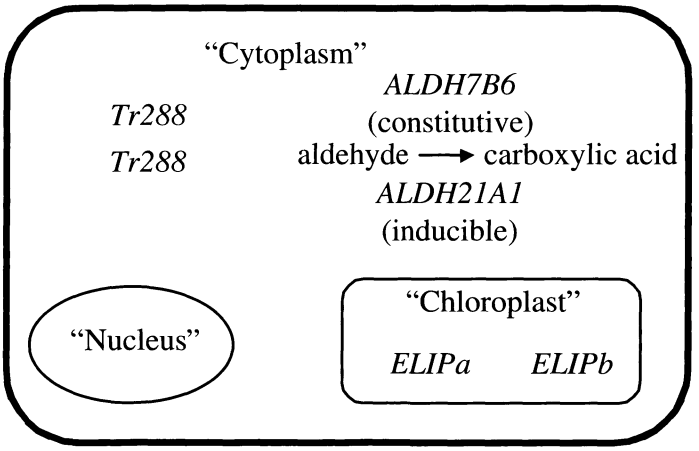

Figure 1: Model depicting the putative sub-cellular localization of the proteins encoded by *Tr288*, *Aldh7B6*, *Aldh21A1*, *Elipa* & *Elipb*.

4.1 The rehydrin Tr288

Tr288 is a rehydrin mRNA which accumulates in the polysomal fraction during drying, is stably maintained in the desiccated state and preferentially translated upon rehydration (Oliver 1991, Scott and Oliver 1994, Velten and Oliver 2001). Tr288 transcript also accumulates when gametophytes are exposed to increasing concentrations of salt and when hydrated gametophytes are subjected to elevated temperatures (Velten and Oliver 2001). Following an advanced BLAST search of Genbank with Tr288 sequences the only significant matches were with several members of the dehydrin (or Late Embryogenesis Abundant (LEA) groupII/LEA D-11) group of plant proteins. The homology was limited to a single copy of the dehydrin "K-segment" (Close 1996,1997) located at the C-terminus of the predicted Tr288 ORF. Tr288 is encoded by a single gene that contains fifteen highly conserved repeated elements that themselves are punctuated by several different inter-repeat sequences (Velten and Oliver 2001). The predicted primary sequence of the Tr288 protein, derived from the ORFs from both a genomic and cDNA clones, reflect the repeated structure of the DNA sequence. The predicted protein has a molecular weight of 68.5Kd which has been confirmed by use of 288 antibody in western blot analysis of moss gametophytic protein extracts (Oliver unpublished data). The protein is highly hydrophilic and glycine rich (19.6%) and contains fifteen copies of a conserved amino-acid sequence motif (called the GPN segment after the highly conserved final three amino acids within the repeat). Computer analysis of the predicted Tr288 protein sequence suggests that the GPN segment and the single dehydrin 'K' segment are likely to form amphipathic (Cornette et. al. 1987), alpha helical regions (Garnier et. al. 1978) (Fig 2).

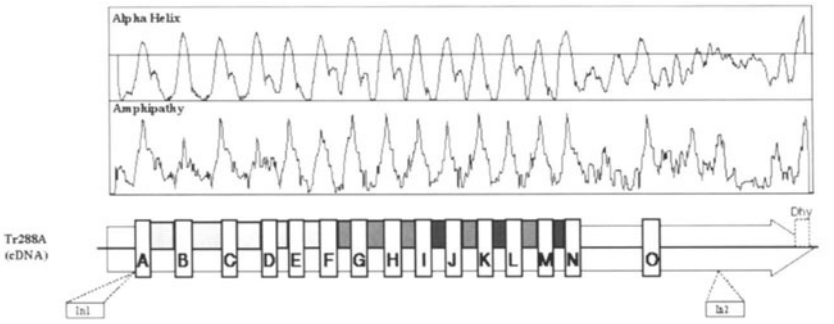

Figure 2: Tr288 cDNA and predicted protein structure. The predicted protein product of the Tr288 cDNA are diagrammatically represented below a computer-generated prediction of protein secondary structure. The fifteen repeated segments are open boxes labeled A-O, each separated by one of three distinct groups of punctuation sequence (shown as light, medium, and dark gray). Introns in the genomic sequence are shown as open boxes below the cDNA diagram.

The repetitive structure and predicted amphipathic helices of the Tr288 gene are reminiscent of some members of the LEA proteins, specifically the D11 (dehydrin) and LEA-D7 gene families (Dure et. al. 1989), as indicated earlier by the observation of a "K-segment" like sequence in Tr288. Such structures are also present in the cold regulated (*COR*) proteins, (Thomashow 1998), and the responsive to ABA (*RAB*) proteins (Xiong et. al. 1999 and Foster and Chua 1999). It is possible that each class of repeated motif proteins share a common structural motif, (i.e., the helical 'core' section of each repeat) linked to differing adjacent residues that serve to define or influence differing functions for each protein group. It has been suggested that the intramolecular redundancy of the various protein motifs argues for a structural or stoichiometric function for the repeated units (Baker et. al. 1988). This argument is even more applicable to the Tr288 gene product due to the large number of repeats (fifteen total). Most discussions of possible LEA/dehydrin protein functions have focused on cellular protection provided prior to, or during, the stress event. In the case of Tr288, the expression (i.e., translation of sequestered mRNA) appears to occur primarily during rehydration and is therefore more likely to be associated with either stabilization during rehydration or repair of cellular components damaged during the drying and/or rehydration processes. There is certainly considerable room for overlap in protection / stabilization verses direct repair activities for these repeated motif proteins (e.g. membrane protection / stabilization / repair). Recently, Koag et al. (2003) reported that a maize dehydrin exhibits lipid binding activity in vitro and that during the binding process the helicity of the protein increases. The implication of this observation is that the dehydrins maybe involved in a membrane stabilization process during desiccation. We have evidence that this is also true for the Tr288 protein (George and Oliver unpublished data) and thus it is possible that the 288 protein in bryophytes functions in membrane stabilization, not during desiccation but during the initial stages of rehydration in order to limit leakage and speed membrane repair.

4. 2 Aldehyde dehydrogenases (ALDHs)

The ALDHs represent a protein superfamily of $NAD(P)^+$-dependent enzymes that oxidize a wide range of range of endogenous and exogenous aliphatic and aromatic aldehydes (EC: 1.2.1.3 & 1.2.1.5) (Vasiliou et al. 1999). Aldehydes are intermediates in a variety of fundamental biochemical pathways and are generated during the metabolism of amino acids, carbohydrates, lipids, biogenic amines, vitamins, and steroids (Yoshida et al. 1998, Vasiliou et al. 2000). Aldehydes can also be generated in response to a variety of environmental stresses that perturb metabolism such as salinity, desiccation, cold, and heat-shock (Barclay and Mckersie 1994). Although common biochemical intermediates, many aldehydes are toxic at physiological concentrations due to their chemical reactivity (Lindahl 1992). In response to UV, desiccation and salinity reactive oxygen species (ROS) are generated that lead to extensive cellular damage including peroxidation of membrane lipids (Smirnoff 1993). Lipid peroxides decompose to aldehydes such as 4-hydroxy-2 (E)-nonenal (4-HNE) and malondialdehyde (MDA) that are postulated to be the toxic second messengers for ROS-induced damage (Esterbauer et

al. 1991). The ALDH superfamily consists of more than 340 distinct genes and the deduced polypeptide sequences share a similar primary structure (Sophos et al. 2000). Phylogenetic comparisons of the eukaryotic sequences have identified 21 unique ALDH protein families based upon sequence homology (Sophos et al. 2000, Wood and Krayesky 2002). Plants are represented in 10 ALDH protein families. 3 protein families are unique to plants (i.e. ALDH10, ALDH11 and ALDH21, and one family (ALDH21) is apparently unique to mosses (Wood and Krayesky 2002). The expression of ALDH-related genes is a common "stress-response" within a number of divergent plant species from mosses to angiosperms.

A number of osmotic stress-inducible ALDH genes have been identified in plants and they comprise the majority of the protein families ALDH7 and ALDH10 (Yoshida et al. 1988). ALDH10 family members encode the drought- and salt-inducible betaine aldehyde dehydrogenase (BALDH) that catalyzes the oxidation of betaine aldehyde to the compatible solute glycine betaine (Rhodes and Hanson 1993). The ability to synthesize and/or accumulate glycine betaine is a ubiquitous adaptation to osmotic stress present in bacteria, animals and plants. However, the biosynthesis and accumulation of glycine betaine is unknown in mosses. The ALDH7 protein family was initially characterized and named as a "turgor-responsive" cDNA in angiosperms (Guerrero et al. 1991). The ALDH7 protein family is further divided into the A sub-family (animal-specific) and B sub-family (plant-specific). A large percentage of subsequently discovered plant ALDH cDNAs are ALDH7 homologues, and the cDNAs have been extensively studied in response to osmotic-stress using RNA gel blot analysis. *ALDH7B1* (26g) from *Pisum sativum* is induced by dehydration but not by heat shock and ABA (Guerrero et al. 1991). *ALDH7B3* (btg-26) from *Brassica napus* is induced by drought, high salinity, low temperature, heat shock, and high concentration of ABA (Stroeher et al. 1995). AtALDH7B4 has been identified in Arabidopsis; however to our knowledge the expression of this gene has not been investigated. Recently, we have cloned an ALDH7 homologue (*Aldh7B6*) from the desiccation-tolerant moss *T. ruralis* (Chen et al. 2002b). ALDH7B6 steady–state mRNA transcript levels are unchanged in response to response to desiccation, NaCl, altered light levels, ABA, & UV-C. We postulate that ALDH7B6 plays an important role in the detoxification of aldehydes generated in response to desiccation-stress and propose that ALDH7 homologues are more properly described as "stress aldehyde" ALDHs. ALDH7 gene products are postulated to be involved in an unknown adaptive metabolic pathway (Guerrero et al. 1991) however, no information on the enzymatic activity or putative function of these proteins has been reported.

Recently, we have also identified a stress-responsive *T. ruralis* cDNA ALDH21A1 that describes a novel eukaryotic aldehyde dehydrogenase protein family (Chen et al. 2002c). The deduced *Tortula* polypeptide is less than 30% identical to prokaryotic and eukaryotic ALDH orthologues (data not shown). According to criteria established by the ALDH Gene Nomenclature Committee (i.e. £40% amino acid identity to known ALDH proteins) (Vasiliou et al. 1999), the *Tortula* sequence describes a new ALDH protein family designated ALDH21. ESTs homologous to ALDH21A1 have been

recently identified in the mosses *P. patens* & *C. purpureus* (AJ Wood, unpublished results). Sequence homology searches (i.e. BlastN, BlastX & BlastP) suggest the ALDH21 protein family is restricted to mosses. Using gametophytes of *T. ruralis*, RNA blot hybridizations were used to analyze expression of *ALDH21A1* in response to desiccation, ABA, UV, and NaCl. *ALDH21A1* steady-state transcript levels increased in response to all treatments and were more abundant within the polysomal mRNA fraction of salt-treated gametophytes. The data suggest that ALDH21A1 plays an important role in the detoxification of aldehydes generated in response to desiccation- and salinity-stress, and we postulate that ALDH21A1 expression represents a unique stress tolerance mechanism.

4.3 Early light-inducible proteins (ELIPs)

ELIPs were first discovered to be transiently expressed during the greening of etiolated plants (Grimm et al. 1989) and are one of the first nuclear-encoded light-inducible proteins detectable within the thylakoid membrane system (Meyer and Kloppstech 1984). ELIPs also accumulate within green tissue in response to a variety of environmental stress factors including high light, UV-B, methyl jasmonate, cold, low oxygen and CO_2 concentration, nutrient starvation, senescence and desiccation (Heddad and Adamska 2002, Zeng et al. 2002). ELIPs contain three transmembrane helices and are members of the extended super family of CABs (Heddad and Adamska 2000). Purified pea ELIPs have been demonstrated to bind both Chl *a* and lutein (Adamska et al. 1999). The membrane insertion of *in vitro* transcribed and translated low molecular mass barley ELIP precursors strictly depended upon chlorophyll *a* but not on chlorophyll *b* nor the xanthophyll zeaxanthin (Adamska et al. 2001). ELIPs are postulated to act as transient pigment-binding proteins during biogenesis or turnover of chlorophyll-binding proteins and protect the chloroplast from light-induced damage that results from the presence of "free" chlorophyll (Lindahl et al. 1997, Adamska et al. 1999). Their function might be extended to include protection against photodamage caused by any situation suboptimal for electron transport through dissipation of excessive light energy (Montane and Kloppstech 2000). ELIPs have been identified and studied in plant tissues that undergo desiccation (i.e. fern spores and leaves of *C. plantagineum*). ELIP-like DSP22 accumulates within *C. plantaginuem* chloroplasts and is postulated to protect the photosystem from desiccation-induced damage (Alamillo and Bartels 2001).

Recently, we have identified two ELIP cDNA orthologues from *T. ruralis*—*Elipa* and *Elipb* (Zeng et al. 2002). *Elipa* steady-state transcript levels increased in response to desiccation (slow-drying), salinity, ABA treatment, low temperature , rapid-drying & rehydration, and rehydration in high light. *Elipb* steady-state transcript levels increased in response ABA treatment and rehydration in high light. Similar to the ELIP dsp22 from the desiccation-tolerant angiosperm *C. plantagineum* (Alamillo and Bartels 2001), *Elipa* is induced in response to both desiccation and ABA. However, *Elipa* and *Elipb* do not respond to environmental factors as a typical angiosperm *Elip* gene. In mature green tissue, ELIPs are expressed in response to high-light and are postulated to protect the chloroplast from high irradiance-induced damage by binding chl *a* and/or zeaxanthin

(Adamska et al. 1999). Neither *Elipa* nor *Elipb* is induced in response to high-irradiance alone (i.e. >500 $\mu Em^{-2}s^{-1}$). In *Tortula*, *Elipa* transcript levels are maximal at 50 $\mu Em^{-2}s^{1}$, remain essentially unchanged up to 1000 $\mu Em^{-2}s^{1}$ and decline at 1500 $\mu Em^{-2}s^{-1}$. With respect to high-irradiance, *Elipa* is constitutively expressed. The photosystems of the desiccation-tolerant mosses *Anomodon viticulosus* and *Racomitrium lanuginosum* rapidly recover upon rewetting (Proctor and Smirnoff 2000) and previous research has emphasized the importance of dynamic non-photochemical quenching (Tuba et al. 1997) and xanthophyll cycle-dependent photoprotection (Gilmore, 1997) in protecting the photosynthetic apparatus of mosses from damage during desiccation. In *T. ruralis*, we hypothesize that salinity, desiccation, rehydration and particularly rehydration in high-light create physiological conditions within the moss chloroplast that mimic chloroplast "greening" and/or exposure to high-irradiance thereby triggering the accumulation of *Elipa* and *Elipb* gene product(s).

4.4 Ubiquitin (Tr 213)

Ubiquitin is a small, highly conserved, 76 amino acid (8.5 kD) protein found in all eukaryotes either as a free monomer or covalently linked to a variety of proteins. Though multifunctional (von Kampen et al. 1996), one of its main known functions is in tagging proteins for selective degradation/turnover by the 26S proteasome (Hasselgren and Fischer, 1997). Ubiquitin is induced by various stresses in plants and animals (Sun and Callis 1997, Binet et al 1991, Christensen et al 1992, Genschik et al 1992, Bond et al 1988, Fornace et al 1989, Fisher et al 1997) but desiccation related effects are poorly understood (Schulz et al 1994, Handke et al 1993). Protein turnover is a normal cellular activity but an increased turnover in response to a stress can be interpreted as the result of excessive protein damage (Ferguson et al 1990) and an attempt to remove damaged proteins from the cell in order to maintain cellular function. This enables us to use ubiquitin as a measure of protein turnover and thus cellular damage/repair in response to desiccation, an important facet of the proposed mechanism of tolerance in *Tortula ruralis*.

Utilizing a partial cDNA clone (Tr213) encoding two copies of the ubiquitin monomer, O'Mahony and Oliver (1999) identified three *Tortula* ubiquitin transcripts representing the ubiquitin gene family; 0.65 Kb, 1.3 Kb and 1.9 Kb, that are present under all conditions. The 0.65 Kb transcript, because of its size, probably encodes a ubiquitin monomer fused at the 3' end to an extension protein, usually a ribosomal protein (Callis et al 1990). This transcript remained at constant levels throughout the drying and rehydration treatments as did the 1.9 Kb transcript. The 1.3 Kb message, however, accumulated in response to rapid and slow drying but only rehydration of rapid dried tissue. This result is consistent with previous results (Scott and Oliver 1994) that demonstrated a ubiquitin transcript accumulates in response rehydration of dried gametophytes. However, the accumulation of the 1.3 Kb ubiquitin transcript during rapid desiccation is in contrast to other rehydrins (proteins whose synthesis is enhanced during rehydration) whose transcripts accumulate during slow drying but do not accumulate during rapid desiccation (Unpublished data and Oliver and Wood 1997). The

accumulation of the 1.3Kb transcript during drying, when it is assumed transcription is inhibited (especially during rapid desiccation), and its storage in the dried state indicates that the transcript is stable under the rigors of water loss. Indeed it appears that all *Tortula* ubiquitin transcripts are stable during desiccation and in the dried state (O'Mahony and Oliver 1999). The accumulation of the 1.3 Kb transcript during drying may indicate that there is an increase in the stabilty of this mRNA, perhaps in rapidly forming mRNPs (Wood and Oliver, 1999), or that the assumption that transcription is inhibited by water loss is an errant one. The significance of the stability and accumulation of ubiquitin transcripts during drying and in the dried state may be linked to the need for the bryophyte to respond rapidly to the cellular damage manifested when gametophytes are rehydrated. The observation that rapid dried gametophytes continue to accumulate the 1.3 Kb transcript following rehydration is consistent with a large body of evidence that rapid desiccation of *Tortula* gametophytes results in greater cellular damage than slow drying (Bewley and Oliver 1992), hence the need for more ubiquitin transcripts.

Western analysis revealed high molecular weight ubiquitin conjugates in *T. ruralis* gametophytes only in slow drying tissue (O'Mahony and Oliver 1999). This, coupled with the transcript analysis, leads us to believe that T. ruralis rapidly and efficiently removes proteins destined for degradation. The appearance of ubiquitin-conjugated proteins in slow drying *T. ruralis* could be interpreted as either a slowing down of the removal of tagged proteins due to water depletion, or the opportunistic removal of damaged proteins that would otherwise be removed during rehydration, thus contributing to a faster recovery. While both interpretations may be valid, the reduced requirement for inducible ubiquitin transcript in rehydrating slow dried *Tortula*, coupled with the appearance of conjugated ubiquitin during slow drying suggests that removal of some dehydration-damaged proteins during desiccation may help to facilitate faster recovery from slow drying. The presence of multiple ubiquitin transcripts, constitutive and inducible, in drying and rehydrating *T. ruralis* suggests that the availability of sufficient ubiquitin monomer in times of critical cell stress may be important for survival of these plants.

6. REFERENCES

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, & Venter, J. C. (1991) Complementary DNA sequencing: Expressed Sequence Tags and Human Genome Project. *Science*. 252, 1651-1656

Adamska, I., Kruse, E., & Kloppstech, K. (2001) Stable insertion of the early light-induced proteins into etioplast membranes requires chlorophyll a. *J. Biol. Chem.* 276, 8582-7

Adamska, I., Roobol-Boza, M., Lindahl, M., & Andersson, B. (1999) Isolation of pigment-binding early light-inducible proteins from pea. *Eur. J. Biochem.* 260, 453-60.

Alamillo, J.M., & Bartels, D. (2001) Effects of desiccation on photosynthesis pigments and the ELIP-like dsp 22 protein complexes in the resurrection plant *Craterostigma plantagineum. Plant Sci.* 160, 1161-70

Alpert P. & Oliver M.J. (2002) Drying without dying. In M. Black & H. W. Pritchard (Eds.) *Desiccation and Survival in Plants: Drying Without Dying* (pp.3-43) CABI Publishing, New York NY

Barclay, K.D. & Mckersie. B. D. (1994) Peroxidation reactions in plant membranes-effects of free fatty acids. *Lipids* 29:877-882

Bewley, J. D. (1979) Physiological aspects of desiccation-tolerance. *Annu. Rev. Plant Physiol.,* 30, 195-238

Bewley, J. D., P. Halmer, J. E. Krochko, & Winner, W. E. (1978) Metabolism of a drought-tolerant and drought-sensitive moss: respiration ATP synthesis and carbohydrate status, In J.H. Crowe and J.S. Clegg (Eds), *Dry Biological Systems*: (pp.185-203). Academic Press, NY

Bewley, J.D. & Oliver, M. J. (1992) Desiccation-tolerance in vegetative plant tissues and seeds: Protein synthesis in relation to desiccation and a potential role for protection and repair mechanisms. In: Osmond, C.B. and Somero, G. (eds.) *Water and Life: A Comparative Analysis of Water Relationships at the Organismic, Cellular and Molecular Levels.* Springer-Verlag, Berlin, pp. 141-160.

Bewley, J.D., Reynolds, T.L., & Oliver, M. J. (1993) Evolving Strategies in the adaptation to desiccation. In: Close, T.J. and Bray, E.A. (eds.) *Plant Responses to Cellular Dehydration During Environmental Stress. Current Topics in Plant Physiology*: American Society of Plant Physiologists Series Vol. 10, pp. 193-201

Binet M.N., Weil J.H., & Tessier, L. H. (1991) Structure and expression of sunflower ubiquitin genes. *Plant Mol. Biol.,* 17, 395-407

Bond U., Agell N., Haas A., Redman K., & Schlesinger, M. (1988) Ubiquitin in stressed chicken embryo fibroblasts. *J Biol Chem* 263, 2384-2388

Callis J., Raasch J., & Vierstra, R. (1990) Ubiquitin extension proteins of Arabidopsis thaliana. Structure, localization, and expression of their promoters in transgenic tobacco. *J. Biol. Chem.* 265, 12486-12493

Chen X., Kanokporn T., Zeng Q., Wilkins T.A., & Wood A.J. (2002a) Characterization of the V-type $H^{(+)}$-ATPase in the resurrection plant *Tortula ruralis*: accumulation and polysomal recruitment of the proteolipid c subunit in response to salt-stress. *J. Expt. Bot.,* 53, 225-232

Chen X., Zeng Q., & Wood A.J. (2002b) *ALDH7B6* encodes a turgor-responsive aldehyde dehydrogenase homologue that is constitutively expressed in *Tortula ruralis* gametophytes. The *Bryologist,* 105, 177-184

Chen X., Zeng Q., & Wood A.J. (2002c) The stress-responsive gene *ALDH21A1* describes a novel eukaryotic aldehyde dehydrogenase protein family. *J. Plant Physiol.*, 159, 677-684

Chen X., & Wood A.J. (2004) The 26S Proteasome of the Resurrection Plant *Tortula ruralis*: Cloning and Characterization of the *TrRPT2* subunit. *Biologia Plant.*, 3, 363-368

Christensen A., Sharrock R., & Quail, P. (1992) Maize polyubiquitin genes: Structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant. Mol. Biol.* 18, 675-689

Diatchenko L., Lau Y-F.C., Cambell A.P., Chenchik A., Moqadam F., Huang B., Lukyanov S., Lukyanov K., Gursskaya N., Sverdlov E.D., & Siebert, P.D. (1996) Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Science USA,* 93, 6025-6030

Duff R.J., Oliver M.J., & Wood A.J. (1999a) Nucleotide sequence of a truncated H3 histone cDNA (accession no. AF093108) from the desiccation-tolerant bryophyte *Tortula ruralis* (PGR99-014). *Plant Physiol.,* 119, 805

Duff R.J., Oliver M.J., & Wood A.J. (1999b) A *Tortula ruralis* cDNA encoding small-subunit ribosomal protein S3a: polysomal retention of transcript in response to desiccation and rehydration. The *Bryologist,* 102, 418-425

Dure, L. III. Crouch, M. Harada, J. Ho, T-HD. & Mundy, J. (1989) Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Mol. Biol.* 12, 475-486

Esterbauer, H., Schaur, R.J., & Zollner, H. (1991) Chemistry and biochemistry of 4-hydroxyenal, malonaldehyde and related aldehydes. *Free Rad. Biol. Med.* 11, 81-128

Fisher E., Zhou M., Mitchell D., Wu X., Omura S., Wang H., Goldberg A. & Ginsberg, H. (1997)The degradation of apolipoprotein B100 is mediated by the ubiquitin-proteasome pathway and involves heat shock protein 70. *J. Biol. Chem.* 272, 20427-20434

Fornace A., Alamo I., Hollander C. & Lamoreaux, E. (1989) Ubiquitin mRNA is a major stress-induced transcript in mammalian cells. *Nucleic Acids Research* 17, 1215-1230

Foster, R. & Chua, N-H. (1999) An Arabidopsis mutant with deregulated ABA gene expression: implications for negative regulator function. *Plant J.* 17, 363-372

Genschik P., Parmentier Y., Durr A., Marbach J., Criqui M.C., Jamet E. & Fleck, J. (1992) Ubiquitin genes are differentially regulated in protoplast-derived cultures of *Nicotiana sylvestris* and in response to various stresses. *Plant Mol. Biol.* 20, 897-910

Gilmore, A.M. (1997) Mechanistic aspects of xanthophyllcycle-dependent photoprotection in higher plant chloroplasts and leaves. *Physiol. Plant.*99, 197-209

Grimm, B., Kruse, E., & Kloppstech, K. (1989) Transiently expressed early light-inducible proteins share transmembrane domains with light-harvesting chlorophyll binding proteins. *Plant Molec. Biol.* 13, 583-93

Guerrero, F.D., Jones, J.T., & Mullet, J. E. (1990) Turgor-responsive gene transcription and RNA levels increase rapidly when pea shoots are wilted. Sequence and expression of three inducible genes. Plant Mol. Biol. 15, 11-16

Handke C., Boyle C., & Wettern, M. (1993) Effect of aging, abiotic and biotic stress upon ubiquitination in young barley plants. *Agnew Bot* 67 120-123

Hasselgren P.O., & Fischer, J. (1997) The ubiquitin-proteasome pathway: Review of a novel intracellular mechanism of muscle protein breakdown during sepsis and other catabolic conditions. *Annals of Surgery* 225, 307-316

Heddad, M., & Adamska, I. (2000) Light stress-regulated two-helix proteins in Arabidopsis thaliana related to the chlorophyll a/b-binding gene family. *Proc. Natl. Acad. Sci. U.S.A* 97, 3741-6

Heddad, M., & Adamska, I. (2002) The evolution of light stress proteins in photosynthetic organisms. *Comp. Funct. Genom.* 3, 504-510

Ingram, J., & Bartels, D. (1996) The molecular basis of dehydration tolerance in plants. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 47, 377-403

Koag, M.C., Fenton R.D., Wilkens S., & Close, T. J. (2003) the binding of maize DHN1 to lipid vesicles. Gain of structure and lipid specificity. *Plant Physiol.* 131, 309-316

Lindahl, R. (1992) Aldehyde dehydrogenases and their role in carcinogenesis *Crit Rev Biochem. Mol. Biol.* 27, 283-33

Lindahl, M., Funk, C., Webster, J., Bingsmark, S., Adamska, I., & Andersson, B. (1997) Expression of ELIPs and PS II-s protein in spinach during acclimative reduction of the photosystem II antenna in response to increased light intensities. *Photo. Res.* 54, 227-36

Machuka J., Bashiardes S., Ruben E., Spooner K., Cuming A., Knight C. D. & Cove D. J. (1999) Sequence analysis of expressed sequence tags from an ABA-treated cDNA library identifies stress response genes in the moss *Physcomitrella patens*. *Plant & Cell Physiol.*, 40, 378-387

Meyer, G., & Kloppstech, K. (1984) A rapidly light-induced chloroplast protein with a high turnover coded for by pea nuclear DNA. *Eur. J. Biochem.* 138, 201-7

Montane, M.H., & Kloppstech, K. (2000) The family of light-harvesting-related proteins (LHCs, ELIPs, HLIPs): was the harvesting of light their primary function? *Gene* 258, 1-8

Oliver M.J., Tuba Z., & Mishler B.D. (2000a) Evolution of desiccation tolerance in land plants. *Plant Ecol.* 151, 85-100

Oliver M.J., Velten J., & Wood A.J. (2000b) Bryophytes as experimental models for the study of environmental stress tolerance: desiccation-tolerance in mosses. *Plant Ecol.*, 151, 73-84

Oliver M.J. (1991) Influence of protoplasmic water loss on the control of protein synthesis in the desiccation-tolerant moss *Tortula ruralis*: Ramifications for a repair-based mechanism of desiccation-tolerance. *Plant Physiol.*, 97, 1501-1511

Oliver, M. J. & Bewley, J. D. (1984a) Desiccation and ultrastructure in bryophytes. *Adv. Bryol.* 2, 91-131

Oliver, M. J. & Bewley, J. D. (1984b) Plant desiccation and protein synthesis: IV. RNA synthesis, stability, and recruitment of RNA into protein synthesis upon rehydration of the desiccation-tolerant moss *Tortula ruralis*. *Plant Physiol.* 74, 21-25

Oliver M.J., & Bewley J.D. (1997) Desiccation-tolerance of plant tissues: a mechanistic overview. *Hort. Rev.*, 18, 171-213

Oliver, M.J., & Wood, A.J. (1997) Desiccation-tolerance of mosses. In T. Koval (Ed), *Stress-inducible Processes in Higher Eukaryotic Cells* (pp. 1-26). Plenum Press, NY.

Oliver M.J., Wood A.J., & O'Mahony P. (1997) How some plants recover from vegetative desiccation: a repair based strategy. *Acta Physiol. Plant.*, 19, 419-425

Oliver M.J., Wood A.J., & O'Mahony P. (1998) "To dryness and beyond"-preparation for the dried state and rehydration in vegetative desiccation-tolerant plants, *Plant Growth Regul.*, 24, 193-201

O'Mahony P.J., & Oliver M.J. (1999) The involvement of ubiquitin in vegetative desiccation tolerance. *Plant Mol. Biol.,* 41, 657-667

Phillips J. R., Oliver M. J. & Bartels D. (2002) Moleulcar genetics of desiccation tolerant systems. In M. Black & H. W. Pritchard (Eds.) *Desiccation and Survival in Plants: Drying Without Dying* (pp.319-341) CABI Publishing, New York NY

Platt, K.A., Oliver, M.J., & Thomson, W. W. (1994) Membranes and organelles of dehydrated *Selaginella* and *Tortula* retain their normal configuration and structural integrity: freeze fracture evidence. *Protoplasma* 178, 57-65

Porembski, S., & Barthlott, W. (2000) Genetic and gneisic outcrops (inselbergs) as centers for diversity of desiccation-tolerant vascular plants. *Plant Ecology,* 151, 19-28

Proctor, M.C.F., & Smirnoff, N. (2000) Rapid recovery of photosystems on rewetting desiccation-tolerant mosses: chlorophyll fluorescence and inhibitor experiments. *J. Expt. Bot* 51, 1695-1704

Proctor, M.C.F. & Tuba, Z. (2002) Poikilohydry and homohydry: anthesis or spectrum of possibilities. *New Phytol.* 156, 327-349

Reski R. S., Reynolds M., Wehe T., Kleberjanke & Kruse S. (1998) Moss (*Physcomitrella patens*) expressed sequence tags include several sequences which are novel for plants. *Bot. Acta,* 111, 143-14

Rhodes, D., & Hanson, A. D. (1993). Quaternary ammonium and tertiary sulphonium compounds in high plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 44, 357-384

Schonbeck, M. W., & Bewley, J. D. (1981). Responses of the moss *Tortula ruralis* to desiccation treatments. II. Variations in desiccation tolerance. *Can. J. Bot.* 59, 2707-2712

Schulz M., Janben M., Knop K., & Schnabl, H. (1994) Stress and age related spots with immunoreactivity to ubiquitin-antibody at protoplast surfaces. *Plant Cell Physiol,* 35, 551-556

Scott II, H.B. & Oliver, M. J. (1994) Accumulation and polysomal recruitment of transcripts in response to desiccation and rehydration of the moss *Tortula ruralis*. *J. Expt. Bot.*, 45, 577-583

Smirnoff, N. (1993) The role of active oxygen in the response of plants to water deficit and desiccation, Tansley Rev. No 52 *New Phytol.* 125, 27:58

Sophos, N.A., Pappa, A., Ziegler, T.L., & Vasiliou, V. (2001) Aldehyde dehydrogenase gene superfamily: the 2000 update. *Chemico-biol. Inter.* 130-132, 323-337

Stroeher, V.L., Boothe, J.G., & Good, A. G. (1995) Molecular cloning and expression of a turgor-responsive gene in *Brassica napus. Plant Mol. Biol.* 27, 541-551

Sun C.W., & Callis, J. (1997) Independent modulation of Arabidopsis thaliana polyubiquitin mRNAs in different organs and in response to environmental changes. *The Plant J.* 11, 1017-1027

Triwitayakorn K, & Wood AJ. (2002) Characterization of two desiccation-stress related cDNAs *TrDr1* and *TrDr2* in the resurrection moss *Tortula ruralis*. *S. Afr. J. Bot.*, 68, 545-548

Thomashow, M.F. (1999) Plant cold acclimation: Freezing tolerance genes and regulatory mechanisms. *Ann. Rev. Plant Phys. Plant Mol. Biol.* 50, 571-599.

Tuba, Z., Proctor, M.C.F., & Csintalan, Z. (1997) Ecophysiological responses of homochlorophyllous and poikilochlorophphylous desiccation tolerant plants. *Plant Growth Reg.* 26, 71

Vasiliou, V., Bairoch, A., Tipton, K. E., & Nebert, D. W. (1999) Eukaryotic aldehyde dehydrogenase (ALDH) genes: human polymorphisms, and recommended nomenclature based on divergent evolution and chromosomal mapping. *Pharmacogen.*, 9, 421-434

Vasiliou, V., Pappa, A., & Petersen, D. R. (2000) Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chemico-biol. Inter.* 129, 1-19

Velten, J., & Oliver, M. J. (2001) Tr288, A rehydrin with a dehydrin twist. *Plant Mol. Biol* 45, 713-722

von Kampen J, Wettern M & Schulz, M. (1996) The ubiquitin system in plants. *Physiologia Plantarum* 97, 618 - 624

Wood A.J., & Duff R.J. (1999) Subunit V (PsaG) of the photosystem I reaction center (accession no. AF157017) from desiccated *Tortula ruralis* (PGR99-140). *Plant Physiol.*, 121, 313

Wood A.J., & Krayesky S.L. (2002) The role of aldehyde dehydrogenases (ALDHs) in abiotic stress tolerance. In A.J. Wood (ed.) Molecular and biochemical responses to the environment (pp. 1-13). Research Signpost. Karachi India

Wood A.J., Duff R.J., & Oliver M.J. (1999) Expressed sequence tags (ESTs) from desiccated *Tortula ruralis* identify a large number of novel plant genes. *Plant & Cell Physiol.*, 40, 361-368

Wood A.J., & Oliver M.J. (1999) Translational control in plant stress: characterization of messenger ribonucleoprotein particles (mRNPs) in desiccated *Tortula ruralis*. *Plant J.*, 18, 359-370

Wood A.J., Duff R.J., & Oliver M.J. (2000a) The translational apparatus of *Tortula ruralis*: polysomal retention of transcripts encoding the ribosomal proteins RPS14, RPS16, and RPL23 in desiccated and rehydrated gametophytes. *J. Expt. Bot.*, 51, 1655-1662

Wood A.J., Duff R.J., Zeng Q., & Oliver M.J. (2000b) Molecular architecture of bryophyte genes: putative polyadenylation signals in cDNA 3'-ends of *Tortula ruralis. The Bryologist*, 103, 44-51

Wood A.J., Oliver M.J., & Cove D.J. (2000c) Frontiers in bryological & lichenological research. I. Bryophytes as model systems. *The Bryologist*, 103, 128-133

Xiong, L. Ishitani, M. & Zhu, J. K. (1999) Interaction of osmotic stress, temperature, and abscisic acid in the regulation of gene expression in Arabidopsis. *Plant J* 17, 363-72.

Yoshida, A., Rzhetsky, A., Hsu, L. C., & Chang, C. (1998) Human aldehyde dehydrogenase gene family. *Eur. J. Biochem.* 251, 549-557

Zeng Q., & Wood A.J. (2000) A cDNA encoding ribosomal protein RPL15 from the desiccation-tolerant bryophyte *Tortula ruralis*: mRNA transcripts are stably maintained in desiccated and rehydrated gametophytes. *Biosci. Biotech. Biochem.*, 64, 2221-2224

Zeng Q., Chen X., & Wood A.J. (2002) Two early light-inducible protein (ELIP) cDNAs from the resurrection plant *Tortula ruralis* are differentially expressed in response to desiccation, rehydration, salinity and high-light. *J. Expt. Bot.,* 53, 1197-1205

EVOLUTION OF THE ORGANELLAR TRANSCRIPTION MACHINERY IN BRYOPHYTES AND VASCULAR PLANTS

YUKIHIRO KABEYA, KOHSUKE SEKINE, NAOKI SATO

Department of Molecular Biology, Saitama University, Shimo-Ohkubo 255, Saitama 338-8570 Japan

Abstract. Plastids are semi-autonomous organelles which possesses their own genetic information. It is generally accepted that plastid arose from a eubacteria-like endosymbiont, closely related to extant cyanobacteria. However, the origin of various DNA-binding proteins and transcription factors in the plastid nucleoid appear complex. Biochemical analysis of these proteins, as well as comparative genome informatics, suggest that proteins of eukaryotic origin replaced most of the original prokaryotic proteins during the evolution of plastids in the lineage of green plants. In addition, nuclear-encoded phage-type RNA polymerase (RPOT), which functions in mitochondria, was recruited, after gene duplication, to the plastids. The creation of nuclear-encoded plastid RNA polymerase (called NEP), as well as the use of NEP in the transcription of certain housekeeping genes of the plastid genome, is a recent event possibly during the evolution of angiosperms. However, each of the two RPOT proteins in a moss *Physcomitrella patens* possesses a potential plastid targeting sequence, which is not translated *in vivo*. Possible roles of the RPOTs in the organelles of the moss are discussed.

1. INTRODUCTION

Plastids and mitochondria are semi-autonomous organelles that possess their own genetic information. As shown in Fig. 1, the plastid DNA is conveniently visualized by fluorescence microscopy with the aid of a DNA-specific fluorochrome, 4',6-diamidino-2-phenylindole (DAPI). The plastid DNA is organized into a particulate structure called 'nucleoid' or 'plastid nucleus' as a result of complex formation with various DNA-binding proteins. Mitochondrial DNA is also organized into mitochondrial nucleoid, although it is not clearly observed in photosynthetic cells, such as a protonemal cell shown in Fig. 1. It is generally accepted that both plastids and mitochondria arose from eubacteria-like endosymbionts, closely related to extant α-proteobacteria and cyanobacteria, respectively (Gray 1992). The entire sequences of the organellar DNA have been determined in many plants such as a liverwort *Marchantia polymorpha* and *Arabidopsis thaliana*, while many bacterial genomes including various species of cyanobacteria have been completely

A.J. Wood et al. (ed.),
New Frontiers in Bryology: Physiology, Molecular Biology and Functional Genomics, 91–110.

sequenced. Recently, the complete nucleotide sequence of the nuclear genome of *A. thaliana* has been determined. These genomic data indicate that many organellar genes were transferred to the host nucleus or lost in various different lineages. In addition, various proteins involved in the maintenance, replication and transcription of the genome have been lost or functionally replaced by other proteins during the evolution of the photosynthetic eukaryotes. Namely, the genome of the cyanobacterium *Synechocystis* sp. PCC 6803 encodes about 40 response regulators, HU protein, and a few transcription factors, such as NtcA and RbcR (Kaneko et al. 1996), whereas none of the homologs of these genes are present in the genomes of both plastid and nucleus of *A. thaliana*. However, in the plastid genomes of chromophytes and rhodophytes, some prokaryotic DNA-binding proteins are encoded such as the HlpA protein, a homolog of bacterial HU protein (Wang and Liu

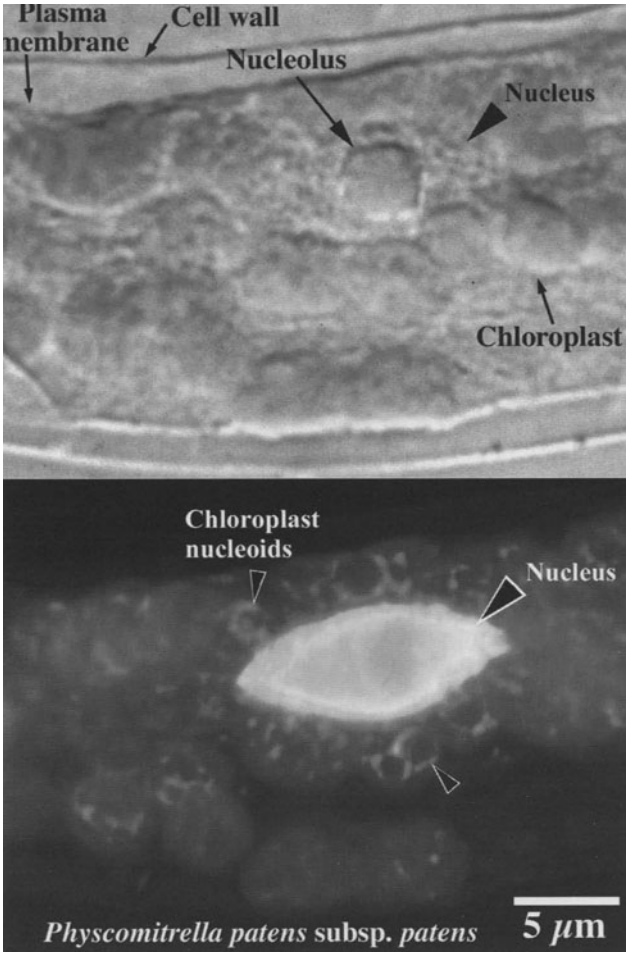

Figure 1. Micrographs of a protonemal cell of Physcomitrella patens.
Upper panel, a Nomarski differential interference image; lower panel, a fluorescence micrograph after staining with DAPI, a fluorochrome that stains DNA.

1991, Kobayashi et al. 2002). Putative transcription factors with similarity to NtcA and others were also encoded in the plastid genome of some rhodophytes or some algae in the rhodophyte lineage such as cryptophytes. In contrast, the DNA-binding proteins that have been characterized biochemically in the plastid nucleoids of vascular plants do not have prokaryotic homologs, such as the CND41 protein, which is known to repress the expression of plastid DNA and has a protease activity (Nakano et al 1997, Murakami et al 2000), and the PEND protein that plays a role in the anchoring of nucleoids to the envelope membrane during the replication of nucleoids (see Fig. 2. Sato et al. 1993, Sato et al. 1998, Sato and Ohta 2001). The structural characteristics of the PEND and CND41 proteins indicate that they probably originate from nuclear transcription factors. Sato et al. (2001) found that sulfite reductase, which is abundant in the plastid nucleoids, possesses DNA-binding activity. Although there are only a few examples at present, it is tempting to assume that the DNA-binding proteins or transcription factors of the plastids of vascular plants are products of importation or re-use of nuclear proteins of eukaryotic origin. In vascular plants, at least two types of plastid RNA polymerases have been identified. One is a prokaryote-type RNA polymerase encoded the plastid genome (called 'PEP') and the other is a phage-type RNA polymerase encoded the nucleus (called 'NEP'; Hedke et al. 1997, Hedke et al. 1999, Hess and Börner 1999, Hedke et al. 2000). It is thought that the gene encoding the NEP arose from a nuclear gene encoding a mitochondrial RNA polymerase by gene duplication (Fig. 2). The NEP can be therefore judged to be of eukaryotic origin. Thus, the plastids lost their prokaryotic DNA-binding proteins and transcription factors, while they acquired eukaryotic DNA-binding proteins and transcription factors from the host nucleus during the evolution in the lineage of green plants.

Sato (2001) put forward a hypothesis on the discontinuous evolution of

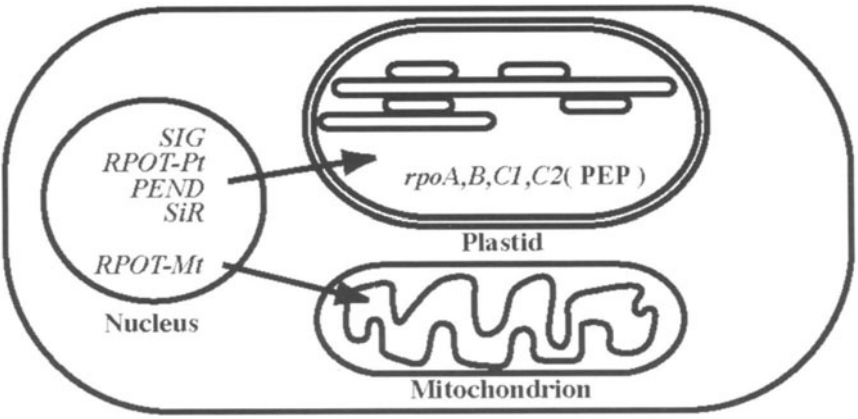

Figure 2. A model of organellar genomic machinery in angiosperm leaf cell. PEP, plastid-encoded RNA polymerase; SIG, gene encoding sigma subunit of plastid PEP; RPOT-Pt, gene encoding phage-type RNA polymerase of plastid; RPOT-Mt, gene encoding phage-type RNA polymerase of mitochondrion; PEND, gene for the plastid envelope DNA-binding protein; SiR, gene for sulfite reductase; rpoA,B,C1,C2, plastid genes encoding subunits of PEP.

plastid genetic machinery: most transcription regulators that had been present in the cyanobacterium-like ancestor were rapidly lost after the primary endosymbiosis, while plastids of land plants acquired transcription factors from the eukaryotic host. This evolutionary change of gene regulation in the plastid genome, which is also the basis of nuclear control of plastid gene expression, must have accompanied various changes in regulatory elements and interactions between components of transcriptional apparatus and regulators. The elucidation of this regulatory shift requires determination of the time or the phyla in which each change occurred. In this respect, bryophytes are candidates, in which eukaryotic proteins were recruited to the plastids. Bryophytes are believed to include early land plants, although the earliest land plants are still in controversy. When the aquatic organisms such as charophytes invaded land surface, the surrounding environment changed drastically. The plastids of land plants must have acquired a number of eukaryotic proteins from the host nucleus in order to adapt themselves to this change. Another change in the plastids during the evolution of vascular plants, is the acquisition of a novel RNA polymerase from the nucleus (NEP), namely, the phage-type RNA polymerase (RPOT). This should have been the driving force of the big changes in the transcription machinery of the plastids, which would have, in turn, affected the evolution of plastids, namely, the ability of the plastids to differentiate. In bryophytes, our knowledge of plastid transcription machinery has been rather limited. Accordingly, it remained unclear if the NEP and other eukaryotic proteins are present in the plastids of bryophytes.

2. GENETIC INFORMATION OF ORGANELLES IN BRYOPHYTES

At present, nucleotide sequences of organellar DNA have been completed in the liverwort *M. polymorpha* (mitochondrion: Oda et al. 1992; plastid: Ohyama et al. 1988), the hornwort *Anthoceros formosae* (plastid: Kugita et al 2001) and *Physcomitrella patens* (plastid: Kobayashi et al. 2002). In addition, some mitochondrial and plastidic genes were identified in many species of bryophytes, *e.g.*, *rbcL* in the plastid genome and the *cox3* in the mitochondrial genome. The *rpoA*, *rpoB, rpoC1* and *rpoC2* genes encoding subunits of the PEP were also identified in various algae and land plants. Interestingly, however, the *rpoA* gene is not present in the plastid genome of *P. patens*, and there are two copies of nuclear *RPOA* genes, each of which encodes a putative RPOA protein that is targeted into the plastid. In the organelles of bryophytes, especially hornworts, a large number of RNA editing sites have been reported (Kugita et al 2001). For example, in the plastid of *A. formosae*, at least 286 editing sites were identified in 44 transcripts. The RNA editing is therefore a possible regulatory point in the post-transcriptional regulation in bryophyte organelles.

The genetic information of organelles in bryophytes have often been used as markers in phylogenetic analysis. The amino acids or nucleic acids sequence of each protein is used in most analyses. However, phylogenetic analysis has also been performed with the presence or absence of group II introns in mitochondrial genes (Qiu et al. 1998, Hashimoto and Sato 2000, Pruchner et al. 2001). The position and the sequence of group II introns are well conserved in many mitochondrial genes. The presence or absence of group II introns in the mitochondrial genes is useful as a

novel marker for phylogenetic analysis of early land plants (Qiu et al. 1998, Nakamura et al. 2002).

3. PLASTID DNA-BINDING PROTEINS IN BRYOPHYTES

3.1. Plastid DNA-binding proteins

In the plastids of vascular plants, various transcription factors have been known: sigma factors transcribe mainly photosynthetic genes together with the PEP (Isono et al. 1997, Fujiwara et al. 2000, Allison 2000, Kanamaru et al. 2001); PTF1 is known as a *trans*-acting factor of the *psbD* light-responsive promoter (Baba et al. 2001); CDF2 is known to regulate the *rrn* gene (Bligny et al. 2000). Additionally, some sequence non-specific DNA-binding proteins are known to constitute the plastid nucleoids. The HC protein was identified by a database search as a basic copper-containing protein (Crevel 1989, Sato 2001), and bifunctional proteins such as the PEND protein, sulfite reductase (abbreviated as 'SiR', Sato et al. 2001) and the CND41 protein have been described. In bryophytes, it is unclear if the transcription factors and DNA-binding proteins that have been studied in higher plants are present. However, nuclear genes encoding sigma factors (Hara et al. 2001a, Hara et al. 2001b) and SiR (Sekine, Kabeya and Sato, unpublished results) were identified in *P. patens*.

2.2. Sigma factors

The prokaryote-type RNA polymerase consists of the core enzyme and a sigma factor. Recently, nuclear genes for putative plastid sigma factors have been identified in algae and land plants (Tanaka et al. 1996, Tanaka et al. 1997, Kestermann et al. 1998, Tan and Troxler 1999, Oikawa et al. 2000). These sigma factor genes (*SIG*) appear to be members of a gene family in each species (Allison 2000). The *SIG* genes in plants appear to originate from the sigma factor genes of the ancestral cyanobacterium-like endosymbiont. In the moss *P. patens*, two genes, each encoding a plastid sigma factor, were identified recently, and named *PpSig1* and *PpSig2*, respectively (Hara et al. 2001a, Hara et al. 2001b). It is known that the *SIG* genes in plants and cyanobacteria are subject to various types of regulation such as tissue-specific (Isono et al. 1997, Tan and Troxler 1999, Beardslee et al. 2002), plastid development-specific (Lahiri and Allison 2000), light-dependent (Kanamaru et al. 1999, Morikawa et al. 1999) and circadian clock-controlled expression (Kanamaru et al. 1999, Morikawa et al. 1999). In addition, different sigma factors are known to play distinct roles under the same light conditions (Oikawa et al. 1998, Tan and Troxler 1999, Beardslee et al. 2002). Both *PpSig1* and *PpSig2* have been shown to be positively regulated by light as are a number of higher plant *SIG* genes. Although studies on the sigma factors in bryophytes are still limited, *SIG* genes are universally present in plants, algae and cyanobacteria, and play important roles in various physiological aspects as listed above.

3.3. DNA-binding proteins in bryophytes

Sulfite reductase (SiR) have been known to be localized to the stroma within the chloroplast (Aketagawa and Tamura 1980). However, in the nucleoids of pea,

SiR was found to be one of the major DNA-binding proteins, which is able to compact DNA as well as isolated nucleoids *in vitro* (Sato et al. 2001, Sekine et al. 2002). Sekine et al. (2002) suggested that SiR reversibly regulates the global transcription activity of plastid nucleoids through changes in DNA compaction. Although SiR is an enzyme catalyzing the reduction of sulfite to sulfide, DNA-binding is its second function in higher plant plastids. Since DNA binding is not known in bacterial SiR, we wondered if SiR from the lower plants or algae has the ability of binding DNA. We performed immunoblot analysis of plastid nucleoids from pea, *P. patens* and a red alga *Cyanidioschyzon merolae* as well as of nucleoids from a cyanobacterium *Anabaena variabilis* strain M3, with antibodies against maize SiR (Sato et al. 2002). Sulfite reductase was detected apparently in the nucleoids of pea and *P. patens*, but not in the nucleoids of *C. merolae* and *A. variabilis*. Sulfite reductase may have acquired the ability to bind DNA during the evolution to mosses. Experiments with green algal nucleoids will be interesting in this respect.

In bacterial nucleoids, the HU protein is known as the major DNA-binding protein. No gene encoding HU has been detected in both plastid and nuclear genomes of *A. thaliana*, but a homolog is present in the plastid genome of a primitive red alga *C. merolae* (Kobayashi et al. 2002) and a cryptophyte *Guillardia theta* (Wang and Liu 1991). The gene for the HU homolog that must have been present in the endosymbiont genome was lost during the evolution of plastids in the chlorophyte lineage (Sato 2001). We performed immunoblot analysis with antibodies against the HU-like protein of *C. merolae*. HU or HU homolog were detected in the nucleoids of *A. variabilis* M3 and *C. merolae* but not in the nucleoids of *P. sativum* and *P. patens* (Sato et al. 2002). A recent study indicated that the HU-like protein from *C. merolae* could also compact DNA as does SiR (Kobayashi et al. 2002, Miyajima and Sato unpublished results). These findings suggest that SiR (eukaryotic major DNA-binding protein in plastid nucleoid) replaced HU (prokaryotic major DNA-binding protein) during the evolution of green plants, although the roles of these two proteins are not identical. Other eukaryotic DNA-binding proteins such as CND41 and PEND have not been found in bryophytes.

4. PHAGE-TYPE RNA POLYMERASES

4.1. Basic characteristics of the phage-type RNA polymerases

Phage-type RNA polymerase (RPOT) is encoded in the nucleus of various eukaryotes, and is known to be targeted to mitochondria or plastids in vascular plants (the one transported to plastids is called NEP, nuclear-encoded RNA polymerase, as described above). In lower land plants and algae, information about organellar transcription machinery is generally limited. In bryophytes, the genes for the prokaryote-type RNA polymerase, encoded in the plastid genome, have been identified in *M. polymorpha* (Ohyama et al. 1988) and *P. patens* (Kobayashi et al. 2002). However, it remained unknown whether the NEP is present or not in bryophytes.

In yeast, the *Rpo41* gene encoding the mitochondrial RNA polymerase was identified by Masters et al. (1987). The disruption experiment indicated that the *Rpo41* gene product is required for the transcription and maintenance of

mitochondrial genome (Masters et al. 1987, Wang and Shadel 1999). It is established that the *Rpo41* gene product plays an important role in transcription and replication (Lecrenier and Foury 2000). In addition, mitochondrial RNA polymerase of *Trypanosoma brucei* is required for transcription and replication (Grams et al. 2002). In higher plants, three types of *RPOT* genes have been identified. The *RpoT;1* gene of *A. thaliana* (Hedke et al. 1997) and the *RpoT1* or *RpoT-A* gene of *N. sylvestris* (Kobayashi et al. 2001, Hedke et al. 2002), with orthologs in the majority of eukaryotic organisms, encode the mitochondrial RNA polymerase. This type of RPOT has been identified in a number of land plants and algae (Cermakian et al. 1996). The *RpoT;3* gene of *A. thaliana* encodes a plastid-targeted RPOT (NEP), which has been identified in only higher plants (Hedke et al. 1997). In addition, the products of the *RpoT;2* gene of *A. thaliana* and the *RpoT2* or *RpoT-B* gene of *N. sylvestris* are dually targeted to mitochondria and plastids (Hedke et al. 2000, Kobayashi et al. 2001, Hedke et al. 2002). In the plastids, the roles of NEP and PEP in transcription are different (Hajdukiewicz et al. 1997, Sakai et al. 1998, Krause et al. 2000). The class I genes (mainly photosystem genes and *rbcL*) are exclusively transcribed by the PEP, whereas the class II genes (genes encoding divergent functions other than photosystems I and II) are transcribed by both PEP and NEP, and the class III genes (mainly housekeeping genes) are transcribed by the NEP. In addition, the promoters recognized by the two RNA polymerases are diverse. There are two classes of promoters, named CT (consensus-type) and NC-II (non consensus-type) promoters, in various plastid genes in higher plants. The CT promoters are recognized by the PEP (Satoh et al. 1999), while the NC-II promoters are recognized by the NEP (Kapoor et al. 1997, Liere and Maliga 1999).

The high similarity of sequences as well as the excellent conservation of

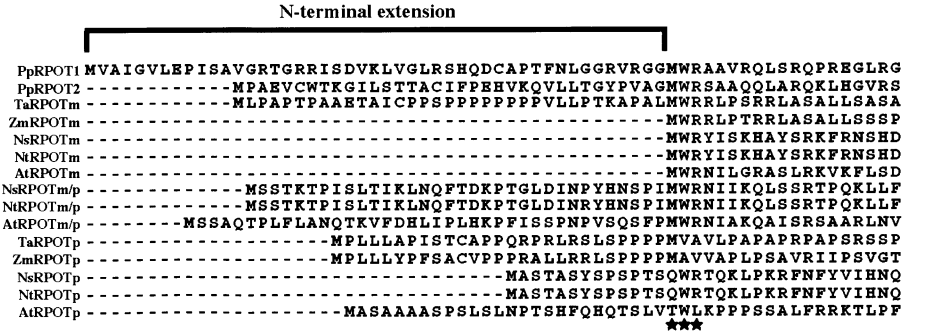

Figure 3. Alignment of N-terminal sequences. The following sequences were obtained ftom the GenBank database: PpRPOT1 and PpRPOT2 (Physcomitrella patens; AB055214 and AB055215); AtRPOTp (Arabidopsis thaliana; Y08463); NsRPOTp (Nicotiana sylvestris; AJ302020); NtRPOTp (Nicotiana tabacum; AJ416575); ZmRPOTp (Zea mays; AF127022); TaRPOTp (Triticum aestivum; U34402); AtRPOTm (Arabidopsis thaliana; Y08137); NsRPOTm (Nicotiana sylvestris; AJ416568); NtRPOTm(Nicotiana tabacum; AJ302018); ZmRPOTm (Zea mays; AJ005343); TaRPOTm (Triticum aestivum; AF091838); AtRPOTm/p (Arabidopsis thaliana; AJ001037); NsRPOTm/p (Nicotiana sylvestris; AJ302019); NtRPOTm/p (Nicotiana tabacum; AJ416569). Asterisks indicate the consrved sequence block.

the positions of introns in the flowering plant *RPOT* genes reported to date suggests that the plastid-type *RPOT* genes are generated from the mitochondrion-type *RPOT* gene by gene duplication. However, we do not know when this gene duplication occurred, because of lack of information on homologs in lower land plants.

4.2. Phage-type RNA polymerase in P. patens

We isolated two cDNAs encoding PpRPOT1 and PpRPOT2 (Kabeya et al. 2002). These cDNAs encoded putative polypeptides of 1,087 and 1,065 residues, respectively. Richter et al. (2002) determined the genomic structure of these genes. The *PpRPOT1* gene consists of 19 exons and 18 introns, while the *PpRPOT2* gene consists of 21 exons and 20 introns. Compared with *A. thaliana*, *PpRPOT1* contains an additional intron in its 5' non-coding region, whereas intron 16 (as of *A. thaliana*) is missing in *PpRPOT1*. The *PpRPOT2* gene contains two additional introns in the 5' non-coding region. The sites of insertion of common introns are precisely conserved between the two *P. patens* genes and the homologs of *A. thaliana*. Both *PpRPOT1* and *PpRPOT2* contain two in-frame AUG codons that might function as initiation sites. The alignment in Fig. 3 shows the N-terminal parts of the two moss sequences and the sequences of known plant RPOTs. The two amino acid residues following the second methionine residue (MWR) are highly conserved in the land plants. However, these residues are not conserved in the homologs in yeast or a primitive red alga (unpublished results), not even in the T7 RNAP. The RPOTs of the mitochondrion-type and the dual targeting-type are both targeted to the mitochondria when translated from the methionine with the asterisk. However, the RPOTs of the dual targeting-type and the plastid-type are targeted to the plastids when translated from the first methionine. Then, what is the localization of RPOTs of *P. patens*? Apparently, the putative transit sequences, each beginning from the first methionine, of the two RPOTs of *P. patens* did not look like the plastid-targeting transit peptides of higher plant RPOTs. The TargetP program (Emanuelsson et al. 2000) produced scores far below the classification threshold (0.04-0.22) for the polypeptides beginning from the first methionine, whereas the polypeptides beginning from the second methionine produced high scores (0.84-0.91) for mitochondrial localization as expected.

Figure 4 shows the phylogenetic tree constructed by the Neighbor-Joining (NJ) method with T7 RNA polymerase as the outgroup. The RPOTs of flowering plants are clearly separated into two groups, each representing plastid- and mitochondrion-localized enzymes. The latter group also includes the dual targeting enzymes. However, the RPOTs from *P. patens* formed a sister group to the cluster of flowering plant RPOTs. In addition, essentially identical results were obtained with the protein and the nucleic acid sequences. High bootstrap supports were obtained in these trees. This phylogenetic analysis suggest clearly that the gene duplication that generated plastid isozymes occurred in the vascular plant lineage after its separation from the moss lineage.

We found a diagnostic region showing distinct characteristic residues specific to mitochondrion-type and plastid-type RNA polymerases of angiosperms as well as to RNA polymerases of yeast, algae, moss and fern (Table 1 and Kabeya et al. 2002). In this region, there are several amino acid residues and deletions that characterized

each group of sequence. The two-residue deletion with angiosperms was not found the PpRPOT1 and PpRPOT2 proteins. In addition, the putative RPOT sequence of *A. capillus-veneris* did not share the deletion. The *Pinus* sequence, though it is not identified as mitochondrial or plastidic enzyme, also retained two amino acids at this site. These results suggest that the gene duplication occurred even later, namely, after the separation of angiosperms from gymnosperms.

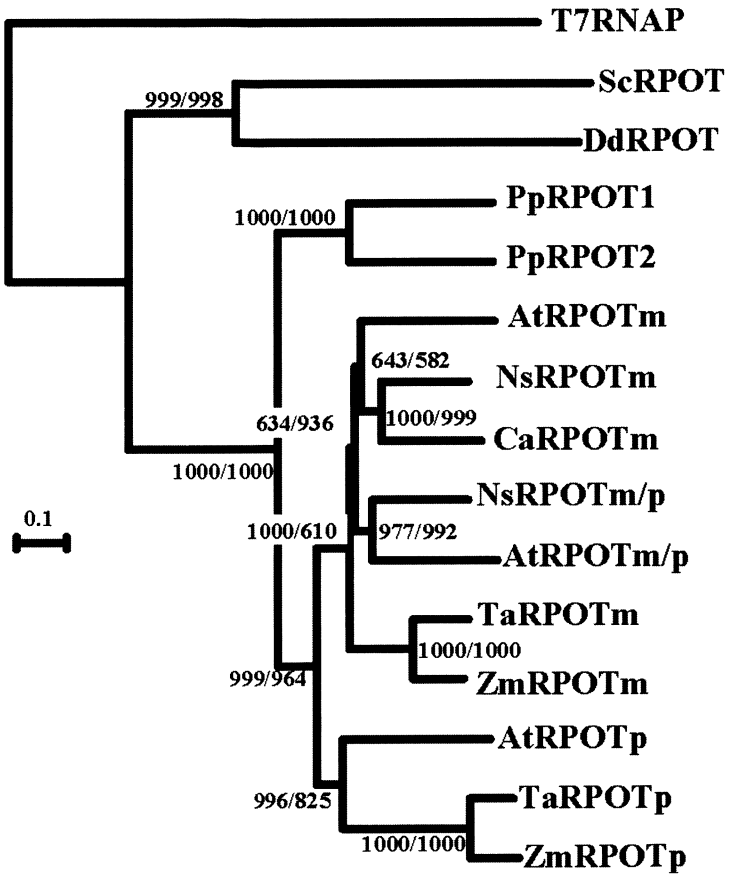

Figure 4. Phylogenetic tree of RNA polymerases. This tree was constructed by the PROTDIST and DNADIST programs using the Neighbor-Joining (NJ) method. The numbers on the branches show bootstrap confidence levels obtained with 1,000 times bootstrap for amino acid / nucleic acid analysis. DdRPOT (Dictyostelium discoideum; GenBank accession AF361480); ScRPOT (Saccharomyces cerevisiae; M17539); CaRPOT (Chenopodium album; Y08067). Other names are explained in Fig. 3.

Table 1. Signature sequence in T7 RNAP and RPOT enzymes. Distinct difference between enzymes of higher plants and others, as well as between mitochondrion- and plastid-type RNAPs were found in this region. Deletion of two amino acid residues was found in higher plant enzymes.

Species	Residues	Sequence
Bacteriophage T7	750-776	M F L G Q F R L Q P T I N T N K D S E I D A H K Q E S
Protists:		
S. cerevisiae	1129-1155	Q V E T N L Q T V F I S D P F A V N P V N A R R Q K A
D. discoideum	822-849	N I R T L E C D F I V V H N D D L L Q V D S N R Q R S
Algae:		
G. theta		S V E T A V Q N I S L L K C D E N G P I N K L K Q R T
P. provasolii		K I N T V L Q T I S V D I H S E A L P V V S A K Q R S
Moss:		
PpRPOT1	960-986	L V K T S L Q V L A L R N L D A D Q P V L V Q R Q K S
PpRPOT2	938-964	L V K T S L Q V L A L R N T D D N H P V L A S R Q R S
Fern:		
A. capillus-veneris		Q V R T S L Q I L A L T D S N D T N M I M V R R Q K S
Gymnosperms:		
P. taeda		L V R T S P Q I W A L R D E T X K V W A I H K K L H S
Angiosperms:		
Mitochondrion-type:		
A. thaliana	851-875	L V K T T L Q V L T L S - - R E T D K V M A R R Q M T
N. sylvestris	876-901	L I K T S L Q I L T L Q - - R E T D K V M V K R Q E T
C. album	863-877	L V K T S L Q V L T L R - - C D T D K V M A K R Q R T
Z. mays	850-874	L I K T S L Q V L T L Q - - R E T D K V M V K R Q K T
T. aestivum	880-904	L I K T S L Q V L T L Q - - R E T D K V M V K R Q R T
Plastid-type:		
A. thaliana	868-892	L I R T S L Q V L A L Q - - R E G N T V D V R K Q R T
Z. mays	827-851	M I R T S L Q C L A L R - - R E G D A I A I Q R Q K A
T. aestivum	824-848	M I R T S L Q C L A L R - - R E G D A I A T Q R Q K A

4.3. Subcellular localization of PpRPOT

In plants such as *A. thaliana, N. sylvestris* and *T. aestivum*, two or three *RPOT* genes have been identified, and the localization of the gene products was found to be in the plastids or mitochondria or both. We previously reported that the two PpRPOTs are localized to only mitochondria (Kabeya et al. 2002). However, Richter et al. (2002) reported that the identical two PpRPOTs are dually targeted to mitochondria and plastids. Namely, both PpRPOTs were targeted to mitochondria when they were translated from the second methionine, whereas they were targeted to plastids when they were translated from the first one. In these experiments, translation was forced to start from the AUG codon just downstream the translation leader sequence taken from tobacco etch virus (using the vector pOL S65C, Peeters et al. 2000). In contrast, we used fusion genes with sGFPS65T (Chiu et al. 1995), which are translated under the natural context with the native 5' leader sequences. The context of the translational initiation site is a likely candidate for the discrepancy.

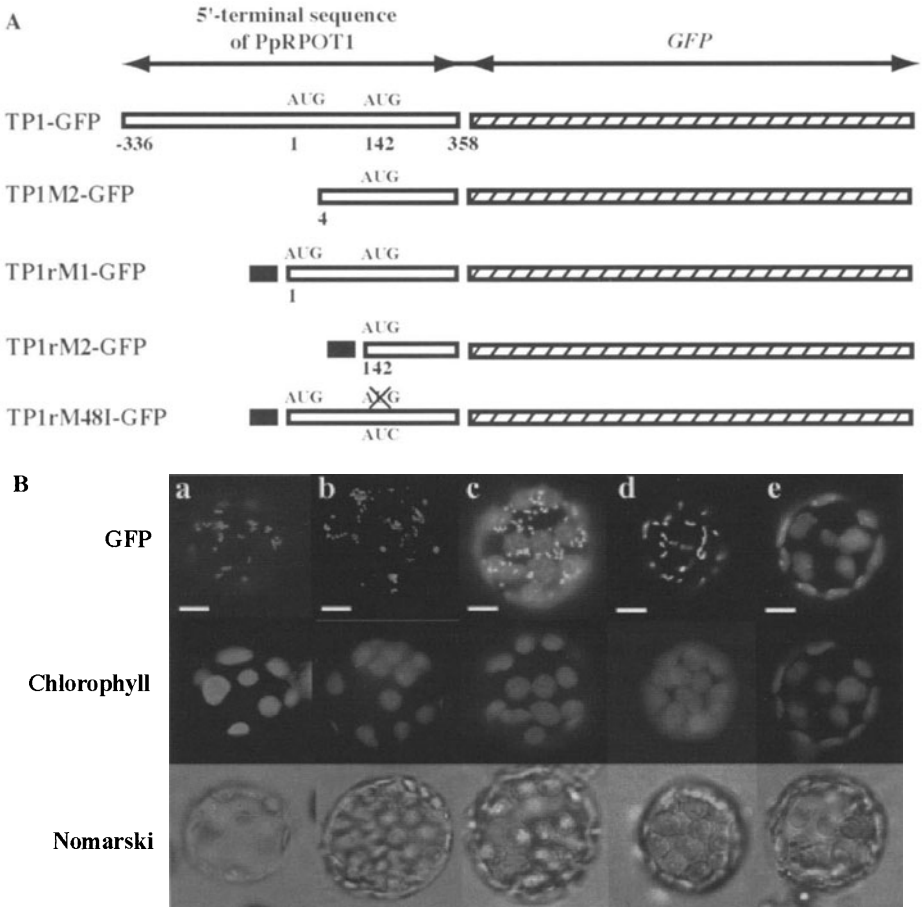

Figure 5. GFP fusion constructs and subcellular localization of GFP-fused PpRPOT1. (A) GFP fusion constructs used for targeting experiments. The 5'-terminal sequences encoding the amino terminus of PpRPOT1 (blank bars) were translationally fused to GFPS65T (hatched bars) in plasmid sGFPS65T. The translation leader sequence from the RbcS 3A gene in pea is indicated as a small filled boxes on the left if added. The length of the fragments is given by nucleic acid residues. The cross on the second AUG codon indicates the mutation introduced at that position to prevent initiation of translation. Bars indicate 5 μm.

To re-examine the subcellular localization of PpPROT1 and PpRPOT2, we prepared new GFP fusion constructs using both natural 5' leader and the translation leader sequence (TL) from the *rbcS* 3A gene of pea (Fig. 5). In addition, in TP1rM48I-GFP, the second initiation codon was mutated, while the translation was forced to start from the first methionine. These constructs were introduced into the moss protoplasts by polyethylene glycol-mediated transformation (Nishiyama et al. 2000, Kabeya et al. 2002). In the experiments with TP1-GFP, TP1M2-GFP and TP1rM2-GFP (Fig. 5B. a, b and d), GFP fluorescence was localized to mitochondria.

The GFP fluorescence of TP1rM48I-GFP was localized to plastids (Fig. 5B. c), and that of TP1rM1-GFP was localized to both mitochondria and plastids (Fig. 5B. e). These results confirmed that the RPOT polypeptide is localized to mitochondria if translated from the second methionine. Forced translatation from the first initiation codon with the TL from pea *rbcS* 3A resulted in the localization to plastids. However, TP1-GFP with the wild-type 5'-upstream sequence was localized to only mitochondria. Therefore, a plausible tentative conclusion is that the first methionine codon is not used as the initiator *in vivo*, although the amino acid sequence beginning from it has a property of plastid targeting sequence. Although we have only the results of new constructs with PpRPOT1, the previous results of Richter et al. (2002) and our results (Kabeya et al. 2002) suggest that the N-terminal extension of PpRPOT2 has also an ability to target to plastids, while the first methionine is again not utilized efficiently *in vivo*. The dual targeting to mitochondria and chloroplasts is now known in various enzymes in angiosperms (Peeters and Small 2001, Watanabe et al. 2001). In all these dually targeted proteins known to date, the use of the first methionine results in targeting to chloroplasts, and the use of the second methionine results in targeting to mitochondria. In many cases, the first methionine is in fact used as the initiation codon *in vivo*. In comparison with these dually targeted enzymes in angiosperms, the case of PpRPOT1 and maybe of PpRPOT2 is very exceptional. The N-terminal extension is a potential plastid targeting sequence, but it is not used *in vivo* as far as the protonematal cells are concerned.

4.4. Effects of tagetitoxin on the organellar transcription in P. patens

The two PpRPOTs were demonstrated to be functional RNA polymerases in a previous study (Kabeya et al. 2002). Here, we investigated effects of tagetitoxin on the transcription in isolated mitochondria and plastids of *P. patens* protonemata (Fig. 4). Tagetitoxin is known to inhibit the transcription by the PEP and bacterial RNA polymerase but not the activity of RPOT (Mathews and Durbin 1990). Transcription activity was measured as an incorporation of [^{3}H]-UTP. In mitochondria, the incorporation of UMP residue into RNA was not inhibited by tagetitoxin, or even stimulated to some extent. No inhibition by tagetitoxin of mitochondrial transcription was just as expected, because mitochondrial transcription is catalyzed solely by phage-type RNA polymerase. In contrast, in the plastids, the incorporation of UMP residue into RNA was perfectly inhibited by the addition of tagetitoxin. In the *in vitro* transcription system using the chloroplasts or the proplastid nucleoids of tobacco BY-2 cells, the residual transcription activity in the presence of tagetitoxin (about 50% of total activity) was ascribed to the NEP (Sakai et al. 1997). In the plastids of *P. patens*, however, there was no trace of tagetitoxin-insensitive transcription activity. Accordingly, the NEP is clearly absent from the plastids of *P. patens* protonemata.

4.5.Implication of the potential plastid targeting sequence

The conflict about the organellar targeting of the two PpRPOTs is now partially resolved. The two RPOT proteins have both two potential initiator methionine codons. The translation beginning from the second methionine codon

results in targeting to mitochondria, whereas the translation from the first one results in targeting to plastids. However, the first methionine codon is not likely to be used as an initiation codon *in vivo* at least in the protonemal cells. This is supported by the targeting experiment with GFP fusions as well as the effect of tagetitoxin on the organellar transcription. The N-terminal extention beginning from the first methionine is therefore a potential plastid targeting sequence, which is normally not utilized. However, the first methionine could be used in some particular types of cells or tissues. We should possibly examine gametophore, bud, sporophyte or primary chloronemal apical cell in the future, but identification of such cells might not be easy.

As described earlier, phylogenetic analysis clearly indicated that the NEP has been created by gene duplication in angiosperms. A supporting evidence for this was obtained by comparison of the promoter regions of NEP-dependent genes. Figure 7 illustrates the gene orders around the *rpoB* and *accD* genes in various plastid genomes. The arrangement around each of these genes in angiosperms is different from that in lower land plants and gymnosperms. Although exact promoter sequences of the *rpoB* and *accD* genes in these plastid genomes are not identified

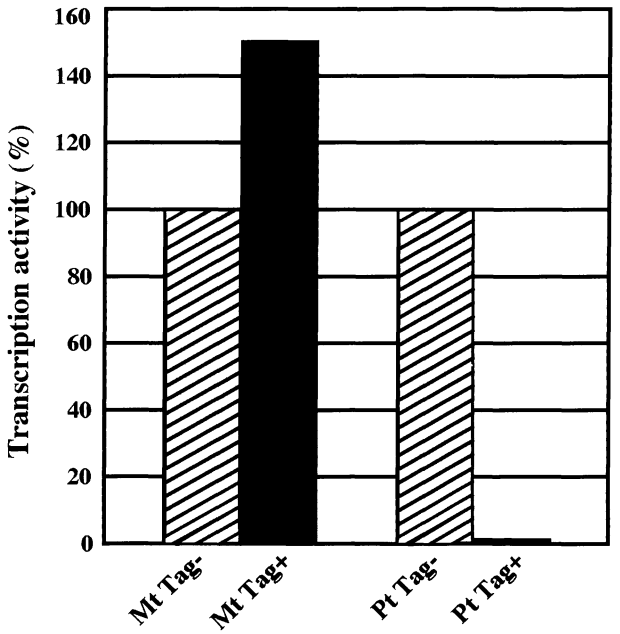

Figure 6. Effects of tagetitoxin on isolated mitochondria and plastids of P. patens. Measurement of the transcription activity was carried out in the presence or absence of tagetitoxin, and [^{3}H] UTP incorporation was determined by liquid scintillation counting. Bars indicate the transcription activity. Black, presence of 10 μM tagetitoxin; hatched, absence of tagetitoxin.

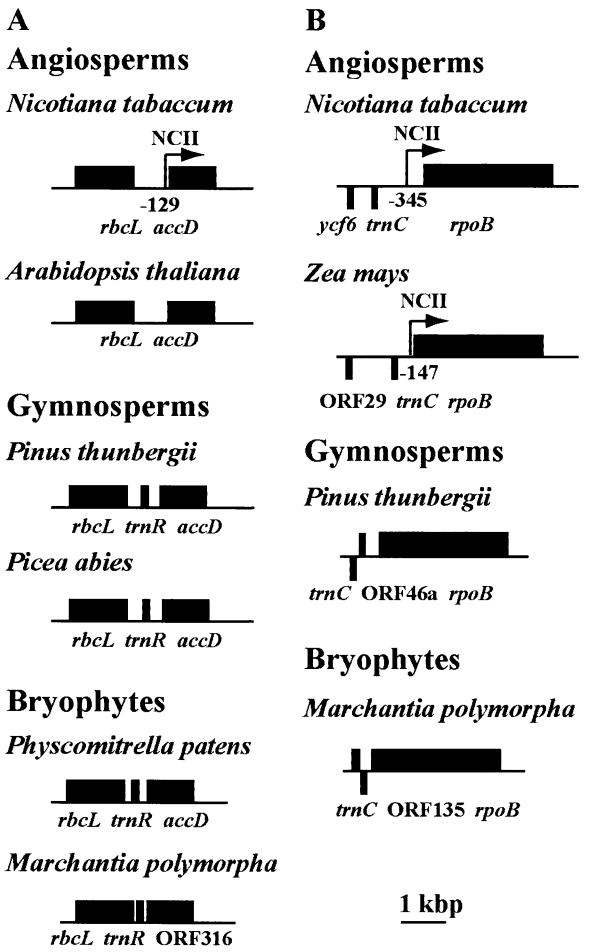

Figure 7. Comparison of gene organization upstream the accD and rpoB genes in various plastid genomes. These two genes are known to be transcribed solely by the NEP in angiosperms. 'NC II' indicates non-consensus-type II promoters, recognized by the NEP, which were identified experimentally. The 5' upstream region of these two genes were subject to reorganization during the evolution of angiosperms. From Kabeya et al. (2002).

experimentally, this comparison suggests that the promoter regions of these genes must have rearranged during the evolution of angiosperms, and that this rearrangement could have created NEP-dependent promoters. It should be noted that, in many monocots, *accD* has been totally lost and another nuclear-encoded acetyl CoA carboxylase consisting of a single multifunctional polypeptide functions in place of the bacteria-type acetyl CoA carboxylase consisting of four subunits. In the land plants other than angiosperms and in algae, the *rpoB* and *accD* genes are, therefore, supposed to be transcribed by the PEP, the product of the *rpoB* operon. It

is tempting to assume that the *rpoB* operon, which encodes subunits of prokaryotic RNA polymerase of plastid, is not totally shut off throughout the life cycle in lower land plants and algae, while in angiosperms, the expression of the *rpoB* operon could be completely shut off in certain stages during the life cycle, possibly in desiccated dormant tissues. In this case, the NEP can switch on the transcription of the *rpoB* operon in the plastid genome, and then, the plastid transcription will be maintained by the PEP. This hypothesis could explain the raison-d'etre of the NEP in angiosperms. The fact that the NEP is mainly expressed in very young developing tissues is consistent with this assumption.

The presence of potential plastid targeting sequence in the two PpRPOTs raises another problem. If the PpRPOTs could be targeted to the plastids by initiation from the first methionine in some specific cells or maybe at a very low rate in protonemal cells, the enzymes would not function as a core of transcription machinery because no specific promoters exist within the moss plastid genome as mentioned above. In yeast mitochondria, it is established that the product of *Rpo41* gene plays a role as a primase in the replication of mitochondrial genome. Therefore, it is possible that the PpRPOTs also function as the primase during the replication of the plastid genome in *P. patens*, if they are eventually imported to the plastids. This could be tested by disruption of RPOT genes in *P. patens*.

Curiously, the RPOT protein has, therefore, repeatedly attempted to invade plastids. The first time was in the moss. The second time was the creation of the NEP in angiosperms. And the third time was the dually targeted RPOT in some angiosperms, probably only in dicots (Fig. 4). In the origin, the mitochondrial RPOT itself was also gained by the eukaryotic cell by horizontal gene transfer. A primitive protozoon *Reclinomonas americana* possesses a bacterial RNA polymerase encoded in the mitochondrial genome (Lang et al. 1997). The mitochondrial genomes of all other eukaryotes that have been studied so far do not encode bacteria-type RNA polymerase. This suggests that this enzyme has been acquired by the common ancestor of most eukaryotic organisms except those in the lineage of *R. americana*. The NEP and the dually targeted RPOT in angiosperms undoubtedly functions as gene-specific polymerases within the plastids, which is dependent on the NC II promoter. However, the plastid genomes of bryophytes are not likely to have this type of promoter. The creation of the NEP in angiosperms must have involved three events, namely, duplication of the *RPOT* gene, gain of the plastid targeting sequence by the RPOT, and gain of the NC II promoter by some genes in the plastid genome. How these complex events occurred concertedly at the beginning of the evolution of angiosperms remains an enigma. In this respect, the case in PpRPOTs suggests that the gain of the plastid targeting sequence preceded the real creation of NEP later in the evolution of plants.

5. CONCLUSION AND PROSPECTS

In the present chapter, we first described essential differences in the components of nucleoids of higher plant plastids and cyanobacteria, and then tried to characterize the bryophyte plastids as an intermediate of these two. The major DNA-binding protein of the cyanobacterial nucleoids is the HU protein, which is also present in the red algal plastids. However, this protein is not present in the land

plants. The major DNA-binding protein of the plastid nucleoids of angiosperms and bryophytes is SiR, a bifunctional enzyme that must have gained the DNA-binding activity during the early stages in the evolution of plants. The sigma factors of the PEP are encoded in the nucleus in all the algae and plants that have been studied. The subunit alpha of the PEP is also encoded in the cell nucleus only in the moss *P. patens*, while it is encoded in the plastid genome of all other land plants studied. The plastids of bryophytes, however, do not have various DNA-binding proteins present in angiosperms, such as PEND, CND41 etc. The mitochondrial RPOT is the origin of the NEP in angiosperms, but NEP is not present in the bryophyte plastids. However, the mitochondrial RPOT proteins in the moss *P. patens* have a potential plastid targeting sequence in their N-terminus, which could be a distant origin of the plastid targeting sequence of angiosperm NEP. These findings support the hypothesis on the discontinuous evolution of plastid genomic machinery (Sato 2001), in which the original endosymbiont had lost most of the prokaryotic components of the nucleoids, whereas the plastids gained many new regulatory proteins from the eukaryotic host during the evolution of the land plants. Here, 'genomic machinery' is intended to include all the components that act on the plastid genome, such as systems of transcription and replication as well as various structural proteins and transcription factors etc. Although replication system is not discussed in this article because of limitation of space, this is also an important focus of study in the forthcoming years.

One driving force of the radical changes in the plastid genomic machinery during the evolution of land plants may be the need for adaptation to the environment of land surface. The environment of the land plants is different in many aspects from that of the algae. High and low temperatures, drought, high light, UV light, are just examples. What is more serious is that the environment changes rather rapidly. The changes in temperature during the day-night cycle as well as the seasonal changes in these environmental conditions are also important in maintaining the life on the land surface. Although these environmental stresses are subject to intense study in agronomy, we are also convinced that better fitness to the harsh environment is attained by the evolution of land plants. One such aspect is the morphogenesis of plastids (and maybe of mitochondria). The chloroplasts of green algae never lose their photosynthetic pigments and ability of photosynthesis. The proplastid in flowering plants might be a form of plastid most adapted to unfavorable environmental conditions. Although many observations have been accumulated in the past on the variable forms of plastids in various different plant tissues and cells, we still do not know exactly the dynamic changes in the components of plastid genomic machinery. The chloroplast development as studied in plant physiology might be the recovery from the completely inactivated state of chloroplast, which was devised during the evolution of the land plants. Sulfite reductase is certainly involved in such repression of plastid genome. The NEP is necessary to re-activate the plastid genome once stalled in the completely repressed state. Many of the topics in physiology can now be placed in and re-organized into a story of plastid evolution.

At present, information on the constituents in the genomic machinery is rather limited, especially in ferns and mosses. However, we have various means to study the evolution of plastid (and eventually mitochondrial) genomic machinery in

land plants. The system for assaying transcription *in vitro* with isolated plastids or nucleoids has a potential to further characterize the difference in plastid transcription machinery between flowering plants and mosses (Fig. 6, and Sakai et al. 1998). Substantial amount of EST (expressed sequence tag) database of *P. patens* have been established in the University of Leeds (UK) and National Institute for Basic Biology (Japan). These databases are useful in identifying new DNA-binding proteins and transcription factors of plastids, as well as various genes dependent on these factors. Experimental identification of nucleoid proteins in *P. patens* and many other plants will also proceed steadily. Advances in the knowledge about plastid genomic machinery in algae, mosses, ferns and flowering plants should shed light on not only the evolution of plastid genomic machinery itself, but also the regulation of plastid gene expression involving crosstalk between the nucleus and the plastids.

6. REFERENCES

Aketagawa, J., & Tamura, G. (1980) Ferredoxin-sulfite reductase drom spinach. *Agric. Biol. Chem.*, 44, 2371-2378

Allison, L., A. (2000) The role of sigma factors in plastid transcription. *Biochimie* 82: 537-548

Baba, K., Nakano, T., Yamagishi, K. and Yoshida, S. (2001) Involvement of a nuclear-encoded basic helix-loop-helix protein in transcription of the light-responsive promoter of *psbD*. *Plant Physiol.*, 125, 595-603

Beardslee, T. A., Roy-Chowdhury, S., Jaiswal, P., Buhot, L., Lerbs-Mache, S, Stern, D. B., & Allison, L. A. (2002) A nucleus-encoded maize protein with sigma factor activity accumulates in mitochondria and chloroplasts. *Plant J.*, 31, 199-209

Bligny, M., Courtois, F., Thaminy, S., Chang, C. C., Lagrange, T., Baruah-Wolff, J., Stern, D., & Lerbs-Mache, S. (2000) Regulation of plastid rDNA transcription by interaction of CDF2 with two different RNA polymerase. *EMBO J.*, 19, 1851-1860

Cermakian, N., Ikeda, M., Cedergren, R , & Gray, M. W. (1996) Sequences homologous to yeast mitochondrial and bacteriophage T3 and T7 RNA polymerases are widespread throughout the eukaryotic lineage. *Nucleic Acids Res.*, 24, 648-654

Chiu, W-I., Niwa, Y., Zeng, W., Hirose, T., Kobayashi, H., & Sheen, J. (1995) Engineered GFP as vital reporter in plants. *Current Biology*, 6, 325-330

Crevel, G., Laine, B., Sautiere, P., & Galleron, C. (1989) Isolation and characterization of DNA-binding proteins from the cyanobacterium Synechococcus sp. PCC 7002 (Agmenellum quadruplicatum) and from spinach chloroplasts. *Biochem. Biophys. Acta.*, 1007, 36-43

Emanuelsson, O., Nielsen, H., Brunak, S., & von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.*, 300, 1005-1016

Fujiwara, M., Nagashima, A., Kanamaru, K., Tanaka, K., & Takahashi, H. (2000) Three new nuclear genes, *sigD*, *sigE*, *sigF*, encoding putative plastid RNA polymerase sigma factors in *Arabidopsis thaliana*. *FEBS Lett.*, 481, 47-52

Grams, J., Morris, J. C., Drew, M. E., Wang, Z., Englund, P. T., & Hajduk, S. L. (2002) A Trypanpsome mitochondrial RNA polymerase is required for transcription and replication. *J. Biol. Chem.*, 277, 16952-16959

Gray, M. W. (1992) The endosymbiont hypothesis revisited. *Int. Rev. Cytol.* 141, 233-357.

Hajdukiewicz, P. T. J., Allison, L. A. and Maliga, P. (1997) The two RNA polymerases encoded by the nuclear and plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J.*, 16, 4041-4048

Hara, K., Sugita, M., & Aoki, S. (2001a) Cloning and characterization of the cDNA for a plastid σ factor from the moss *Physcomitrella patens*. *Biochem. Biophys. Acta.*, 1517, 302-306

Hara, K., Morita, M., Takahashi, R., Sugita, M., Kato, S., & Aoki, S. (2001b) Characterization of two genes, *Sig1* and *Sig2*, encoding distinct plastid σ factors in the moss *Physcomitrella patens*: phylogenetic relationships to plastid σ factors in higher plants. *FEBS Lett.*, 499, 87-91

Hashimoto, K., & Sato, N. (2000) Characterization of the mitochondrial *nad7* gene *in Physcomitrella patens*: Similarity with angiosperm *nad7* genes. *Plant Science*, 160, 807-815

Hedtke, B., Börner, T., & Weihe, A. (1997) Mitochondrial and chloroplast phage-type RNA polymerases in *Arabidopsis. Science,* 277, 809-811

Hedtke, B., Meixner, M., Gillandt, S., Richter, E., Börner, T., & Weihe, A. (1999) Green fluorescent protein as a marker to investigate targeting of organellae RNA polymerases of higher plants in vivo. *Plant J.*, 17, 557-561.

Hedtke, B., Börner, T., & Weihe, A. (2000) One RNA polymerase serving two genomes. *EMBO Rep.,* 1, 435-440.

Hedtke, B., Legen, J., Weihe, A., Herrman, R. G., & Börner, T. (2002) Six active phage-type RNA polymerase genes in *Nicotiana tabacum. Plant J.*, 30, 625-637

Hess, W. R., & Börner, T. (1999) Organellar RNA polymerases of higher plants. *Int. Rev. Cytol.*, 190, 1-59.

Isono, K., Shimizu, M., Yoshimoto, K., Niwa, Y., Satoh, K., Yokota, A., & Kobayashi, H. (1997) Leaf-specifically expressed genes for polypeptides destined for chloroplasts with domains for σ^{70} factors of bacterial RNA polymerases in *Arabidopsis thaliana. Proc. Natl. Acad. Sci. U. S. A.*, 94, 14948-14953

Kabeya, Y., Hashimoto, K., & Sato, N. (2002) Identification and characterization of two phage-type RNA polymerase cDNAs in the moss *physcomitrella patens*: Implication of recent evolution of nuclear-encoded RNA polymerase of plastids in plants. *Plant Cell Physiol.*, 43, 245-255

Kanamaru, K., Fujiwara, M., Seki, M., Katagiri, T., Nakamura, M., Mochizuki, N., shinozaki K., Tanaka, K., & Takahashi, H. (1999) Plastidic RNA polymerase σ factors in *Arabidopsis. Plant Cell Physiol.*, 41, 832-842

Kanamaru, K., Nagashima, A., Fujiwara, M., Shimada, H., Shirano, Y., Nakabayashi, K., Shibata, D., Tanaka, K., & Takahashi, H. (2001) An Arabidopsis sigma factor (SIG2)-dependent expression of plastid-encoded tRNAs in chloroplasts. *Plant Cell Physiol.*, 42, 1034-1043

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., & Tabata S. (1996) Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.*, 3, 109-136

Kapoor, S., Suzuki, J. Y., & Sugiura, M. (1997) Identification and functional significance of a new class of non-consensus-type plastid promoters. *Plant J.*, 11, 327-337.

Kestermann, M., Neukirchen, S., Kloppstech, K., & Link, G. (1998) Sequence and expression characteristics of a nuclear-encoded chloroplast sigma factor from mustard (*Sinapis alba*). *Nucleic Acids Res.*, 26, 2747-2753

Kobayashi, T., Takahara, M., Miyagishima, S., Kuroiwa, H., Sasaki, N., Ohta, N., Matsuzaki, M. and Kuroiwa, T. (2002) Detection and localization of a chloroplast-encoded HU-like protein that organizes chloroplast nucleoids. *Plant Cell*, 14, 1579-1589

Kobayashi, Y., Dokiya, Y., Sugiura, M., Niwa, Y., & Sugita, M. (2001a) Genomic organization and organ-specific expression of a nuclear gene encoding phage-type RNA polymerase in *Nicotiana sylvestris*. *Gene*, 279, 33-40

Kobayashi, Y., Dokiya, Y., & Sugita, M. (2001b) Dual targeting phage-type RNA polymerase to both mitochondria and plastids is due to alternative translation initiation in single transcripts. *Biochem. Biophys. Res. Commun.*, 289, 1106-1113

Kobayashi, Y., Sugiura, C., & Sugita, M. (2002) The moss Physcomitrella patens chloroplast rpoA gene is present in the nuclear genome. *MOSS2002 Abstract Book*

Krause, M., Maier, R. M., Kofer, W., Krupinska, K., & Herrmann, R. G. (2000) Disruption of plastid-encoded RNA polymerase genes in tobacco: expression of only a distinct set of genes is not based on selective transcription of the plastid chromosome. *Mol. Gen. Genet.*, 263, 1022-1030

Kugita, M., Fujikawa, A., Kaneko, A., Yamamoto, Y., Takeya, Y., & Yoshinaga, K. (2001) Complete sequence of the hornwort (*Anthoceros formosae*) chloroplast genetics : phylogenety of early land plants inferred from chloroplast genomes. *MOSS2001 Abstract Book*

Lahiri, S. D., & Allison, L. A. (2000) Complementaly expression of two plastid-localized sigma-likefactors in maize. *Plant Physiol.*, 123, 883-894

Lang, B. F., Burger, G., O'Kelly, C. J., Cedergren, R., Golding, G. B., Lemieux, C., Sankoff, D., Turmel, M. and Gray, M. W. (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature*, 387, 493-497.

Lecrenier, N., & Froury, F. (2000) New features of mitochondrial DNA replication system in yeast and man. *Gene*, 246, 37-48

Liere, K., & Maliga, P. (1999) In vitro characterization of the tobacco *rpoB* promoter reveals a core sequence motif conserved between phage-type plastid and plant mitochondrial promoters. *EMBO J.*, 18, 249-257

Masters, B. S., Stohl, L. L., & Clyton D. A. (1987) Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. *Cell*, 51, 89-99

Mathews, D. E., & Durbin, R. D. (1990) Tagetitoxin inhibits RNA synthesis directed by RNA polymerases from chloroplasts and *Escherichia coli*. *J. Biol. Chem.*, 265, 493-498

Morikawa, K., Ito, S., Tsunoyama, Y., Nakahira, Y., Shirano, T., & Toyoshima, Y. (1999) Circadian-regulated expression of a nuclear-encoded plastid sigma factor gene (sigA) in wheat seedling. *FEBS Lett.*, 451, 275-278

Murakami, S., Kondo, Y., Nakano, T., & Sato, F. (2000) Protease activity of CND41, a chloroplast nucleoid DNA-binding protein, isolated from cultured tobacco cells. *FEBS Lett.*, 468, 15-18

Nakamura, M., Kabeya, Y., & Sato, N. (2002) Phylogenetic analysis of land plants using introns within the *nad7* gene. *Plant Cell Physiol.*, 43, s123

Nakano, T., Murakami, S., Shoji, T., Yoshida, S., Yamada, Y., & Sato, F. (1997) A novel protein with DNA binding activity from tobacco chloroplast nucleoids. *Plant Cell*, 9, 1673-1682

Nishiyama, T., Hiwatashi, Y., Sakakibara, I., Kato, M., & Hasebe, M. (2000) Tagged mutagenesis and gene-trap in the moss, *Physcomitrella patens* by shuttle mutagenesis. *DNA Res.*, 7, 9-17

Oda, K., Yamamoto, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K., Kanegae, T., Ogura, Y., & Kohchi, T., et al. (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive from of plant mitochondrial genome. *J. Mol. Evol.*, 223,1-7

Ohyama, K., Fukuzawa, H., Kohchi, T., Sano, T., Sano, S., Shirai, H., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., & et al. (1988) Structure and organization of *Marchantia polymorpha* chloroplast genome. I. Cloning and gene identification. *J. Mol. Biol.*, 203, 281-298

Oikawa, K., Fujiwara, M., Nakazato, E., Tanaka, K., & Takahashi, H. (2000) Characterization of two plastid σ factors, SigA and SigB, that mainly function in matured chloroplasts in *Nicotiana tabacum*. *Gene*, 261, 221-228

Peeters, N. M., Chapron, A., Giritch, A., Grandjean, O., Lancelin, D., Lhomme, T., Vivrel, A., & Small, I. (2000) Duplication and quadruplication of Arabidopsis thaliana cysteinyl- and Asparaginyl-tRNA synthetase genes of organellar origin. *J. Mol. Evol.*, 50, 413-423

Peeters, N., & Small, I. (2001) Dual targeting to mitochondria and chloroplasts. *Biochem. Biophys. Acta.*, 1541, 54-63

Pruchner, D., Nassal, B., Schindler, M., & Knoop, V. (2001) Mosses share mitochondrial group II introns with flowering plants, not with liverworts. *Mol. Genet. Genomics*, 266, 608-613

Qiu, Y. L., Cho, Y., Cox, J. C., & Palmer, J. D. (1998) The gain of three mitochondrial introns identifies liverworts as the earliest land plants. *Natur*e, 394, 671-674

Richter, U., Kiessling, J., Hedtke, B., Decker, E., Reski, R., Börner, T., & Weihe, A. (2002) Two *RpoT* genes of *Physcomitrella patens* encode phage-type RNA polymerases with dual targeting to mitochondria and plastids. *Gene*, 290, 95-105

Sakai, A., Saito, C., Inada, N., & Kuroiwa, T. (1997) Transcriptional activities of the chloroplast-nuclei and proplastid-nuclei isolated from tobacco exhibit different sensitivities to tagetitoxin: implicatikon of the presence of distinct RNA polymerase. *Plant Cell Physiol.*, 39, 923-934

Sakai, A., Suzuki, T., Miyazawa, Y., & Kuroiwa, T. (1998) Simultaneous isolation of cell-nuclei, plastid-nuclei and Mitochondrial-nuclei from cultured tobacco cells; comparative analysis of their transcriptional activities in invitro. *Plant Science*, 133, 17-31

Sato, N., Albrieux, C., Joyard, J., Douce, R., & Kuroiwa, T. (1993) Detection and characterization of a plastid envelope DNA binding protein which may anchor plastids nucleoids. *EMBO J.*, 12, 555-561

Sato, N., Oshima, K., Watanabe, A., Nishiyama, Y., Joyard, L., & Douce, R. (1998) molecular characterization of the PEND protein, a novel bZIP protein in the envelope membrane that is the site of nucleoid replication in developing plastids. *Plant Cell*, 10, 859-872

Sato, N., & Ohta, N. (2001) DNA-binding specificity and dimerization of the DNA-binding domain of the PEND protein in the chloroplast envelope membrane. *Nucleic Acids Res.*, 29, 2244-2250

Sato, N., Nakayama, M., & Hase, T. (2001) The 70-kDa major DNA-compacting protein of the chloroplast nucleoid is sulfite reductase. *FEBS Lett.*, 487, 347-350

Sato, N., Sekine, K., Kabeya, Y., Ehira, S., Onuma, M., & Ohta, N. (2002) Discontinuous evolution of plastid genomic machinery: Radical replacement of major DNA-binding proteins. In "(Endo)Symbiosis and Eukaryotic Organells" (Sugiura, M. ed.), Logos Verlag Berlin (in press)

Satoh, J., Baba, K., Nakahira, Y., Tsunoyama, Y., Shiina, T., & Toyoshima, Y. (1999) Developmental

stage-specific multi-subunit plastid RNA polymerases (PEP) in wheat. *Plant J.*, 18, 407-415
Sekine, K., Hase, T., & Sato, N. (2002) Reversible DNA conpaction by sulfite reductase regulates transcriptional activity of chloroplast nucleoids. *J. Biol. Chem.*, 277, 24399-24404
Tan, S., & Troxler, R. (1999) Characterization of two chloroplast RNA polymerse sigma factors from Zea mays: photoregulation and differential expression. *Proc. Natl. Acad. Sci. U. S. A.*, 96, 5316-5321
Tanaka, K., Oikawa, K., Ohta, N., Kuroiwa, H., Kuroiwa, T., & Takahashi, H. (1996) Nuclear encoding of a chloroplast RNA polymerase sigma subunit in a red alga. *Science*, 272, 1932-1935
Tanaka, K., Tozawa, Y., Mochizuki, N., Shinozaki, K., Nagatani, A., Wakasa, K., & Takahashi, H. (1997) Characterization of three cDNA species encoding plastid RNA polymerase sigma factors in Arabidopsis thaliana: Evidence for the sigma factor heterogeneity in higher plant plastids. *FEBS Lett.*, 413, 309-313
Wang, S., & Liu, X. (1991) The plastid genome of *Cryptomonas* encodes an hsp 70-like protein, a histone-like protein, and an acyl carrier protein. *Proc. Natl. Acad. Sci. U. S. A.*, 88, 10783-10787
Wang, Y., & Shadel, G. (1999) Stability of the mitochondrial genome requires an amino-terminal domain of yeast mitochondrial RNA polymerse. *Proc. Natl. Acad. Sci. U. S. A.*, 96, 8046-8051
Watanabe, N., Che, F. S., Iwano, M., Takayama, S., Yoshida, S., & Isogai, A. (2001) Dual targeting of spinach protoporphyrinogen oxidase II to mitochondria and chloroplasts by alternative use of two in-frame initiation codons. *J. Biol. Chem.*, 276, 20474-20481

GENE TAGGING, GENE- AND ENHANCER-TRAPPING, AND FULL-LENGTH cDNA OVEREXPRESSION IN *PHYSCOMITRELLA PATENS*

TOMOMICHI FUJITA, TOMOAKI NISHIYAMA, YUJI HIWATASHI, MITSUYASU HASEBE

National Institute for Basic Biology, Okazaki 444-8585, JAPAN

Abstract: Physcomitrella (*Physcomitrella patens*) is a useful model in plant biology, because of its simpler morphology and development as compared with other multicellular model plants and its suitability for gene targeting. Physcomitrella has the highest homologous recombination rate of any green plant examined to date, which can be utilized for the highly reliable gene targeting technique. Recent technical advances in genome biology make Physcomitrella even more useful. A shuttle mutagenesis technique based on homologous recombination was used to generate a large number of insertional mutant lines, as well as gene-trap and enhance-trap lines. More than 12,000 gene-trap lines were collected and the trapped genes, which showed cell- or tissue-specific expression patterns, were successfully identified. High quality full-length cDNA resources were established from the haploid generation of Physcomitrella; these resources are useful for functional analyses of genes and their products. More than 40,000 cDNAs have been sequenced from the both ends, and over 100,000 ESTs are publicly available. Sequence comparisons between Arabidopsis (*Arabidopsis thaliana*) and Physcomitrella revealed that two thirds of Arabidopsis genes had homologs in Physcomitrella. Physcomitrella should therefore be useful for analyzing the detailed function of genes originally isolated in other plant models. We identified 300 candidates for moss transcripts that are absent in vascular plants, and which likely represent new genes in land plants. Analysis of these genes will provide further new genetic resources. An over-expression screening system for full-length cDNA has been established to screen genes involved in cell polarity initiation and for maintenance during protoplast regeneration.

1. INTRODUCTION

The sequences of the entire Arabidopsis (*Arabidopsis thaliana*) and rice (*Oryza sativa*) genomes have been determined, and genome projects for several other plant species are in progress. In this genomic (or post-genomic) era, defining the functions of genes and their products is a key area of study for plant biologists. In recent decades, molecular genetic approaches have proven to be the most powerful techniques available for revealing gene function. The loss-of-function approach, derived from gene disruption by chemical mutagenesis or insertional mutagenesis, has been widely applied in model plants such as Arabidopsis.

Recently, the moss *Physcomitrella* (*Physcomitrella patens*) has emerged as a new model plant for functional genomic studies due, in particular, to its

A.J. Wood et al. (ed.),
New Frontiers in Bryology: Physiology, Molecular Biology and Functional Genomics, 111–132.

amenability to gene targeting methods, which utilize its efficient homologous recombination, a feature unique among plants (Schaefer 2002). However, the use of forward genetic approaches is not as straightforward in *Physcomitrella* as it is with other model plants. Positional cloning to identify a mutated gene following chemical mutagenesis is not feasible because of the relatively poor *Physcomitrella* genome information. Additionally, *Agrobacterium*- or transposon-mediated gene-tagging methods, which enable isolation of genes based on loss-of-function phenotypes in organisms for which genetic maps are unavailable, has not been reported for *Physcomitrella*. Preliminary tests of several *Agrobacterium* strains for use in transformation of *Physcomitrella* have not been successful (Cove et al. 1997), and the Ac/Ds transposon system, established in maize and also used in Arabidopsis, is not yet applicable to *Physcomitrella* (Knight and Cove 1993). Shuttle mutagenesis is the only method available at present to establish tagged mutants in *Physcomitrella*. This method has been used to generate a large number of tagged lines with interesting phenotypes, possibly caused by loss-of-function mutations (Egener et al. 2002, Hiwatashi et al. 2001, Nishiyama et al. 2002).

Although loss-of-function approaches are indispensable to certain aspects of plant science, it is also clear that not all genes can be identified by the saturation mutagenesis of such methods. Some genes are functionally redundant, and mutation of these genes is unlikely to result in an easily recognizable phenotype. Other genes may function at varying stages of development, and mutations in these genes could result in a highly pleiotropic phenotype, which could mask their roles in specific pathways (Springer 2000). Gene-trap and enhancer-trap systems in plants and animals are powerful tools for finding novel genes and unidentified regulatory elements, based on the expression of a reporter gene (Bellen 1999, Springer 2000). Shuttle mutagenesis is effective for the delivery of gene-trap or enhancer-trap elements into the *Physcomitrella* genome, and a large collection of gene-trap and enhancer-trap lines have been generated (Hiwatashi et al. 2001, Nishiyama et al. 2002). We describe here the successful isolation of trapped genes from gene-trap lines.

High-quality full-length cDNA libraries were generated from haploid-generation material of *Physcomitrella* that had been grown with or without phytohormone treatments, and EST sequences from these libraries are available in public DNA databases. The approximately 13,000 independent genes represented by these ESTs are useful resources for Physcomitrella functional genomics (Mullins and Kang 2001). One valuable application of the full-length cDNA sequence information is its use in gain-of-function analysis *via* overexpression of the genes. We report here a preliminary screen for polarity mutants using transient overexpression of a full-length cDNA. These new resources, such as trap lines and the collection of full-length cDNA clones, render *Physcomitrella* one of the most suitable model plants in the post-genomic era.

2. TAGGED MUTAGENESIS, GENE-TRAP, AND ENHANCER-TRAP SYSTEMS MEDIATED BY SHUTTLE MUTAGENESIS

2.1. Tagged Mutant Lines Generated by Shuttle Mutagenesis

Shuttle mutagenesis was developed in the budding yeast *Saccharomyces cerevisiae* (Siefert et al. 1986a, b) by exploiting its high homologous recombination rate. This method is also well-suited to other organisms with similarly high homologous recombination rates, such as *Physcomitrella*. The procedure, shown in Fig. 1a, involves three steps: (1) generation of a host genomic DNA library; (2) insertion of a mini-transposon into the host genomic DNA fragment in *Escherichia coli*; and (3) transformation of host cells and replacement of the genomic locus with the genomic DNA, tagged with the mini-transposon, using homologous recombination. The tag can be inserted in various positions in the host genome, and the sequence flanking the tag can be isolated by a genomic-sequence-based approach, such as PCR, or by hybridization screening of a genomic DNA library prepared from the tagged lines (Nishiyama et al. 2000). Shuttle mutagenesis has been applied in Physcomitrella, with some modifications of vector constructs. An original mini-transposon element derived from Tn3 was modified to generate mTn-nptII, which includes a cauliflower mosaic virus (CaMV) 35S promoter that drives an nptII cassette, allowing the selection of transformed Physcomitrella lines on medium containing antibiotic G418 (Fig. 1b) (Schaefer et al. 1991). Two other mini-transposons, HI-GT and HI-ET, were used to generate tagged lines; these fragments also functioned as gene-trap and enhancer-trap elements, respectively (Fig. 1c and d, see details in Section 2.2). Egener et al. (2002) established a system based on another transposon, Tn1000, and a nos-promoter-driven nptII cassette.

A *Physcomitrella* genomic library was mutagenized with the modified transposon using a series of *E. coli* strains Nishiyama et al. 2000). Plasmids containing transposon-tagged Physcomitrella genomic DNA fragments were extracted from *E. coli*, linearized with a restriction enzyme, and used to transform Physcomitrella protoplasts with the polyethylene glycol (PEG)-mediated method. Linearization was employed to avoid multiple copies of the insert at an insertion site. Construction of the genomic library was difficult, depending on the trans-illuminator used, possibly due to exposure of the genomic DNA fragments to UV light, which decreases the primary cloning efficiency. Staining of the genomic DNA fragments with Crystal Violet (Sigma-Aldrich, St. Louis, MO) and observation under visible light may result in more efficient and reproducible cloning than staining with ethidium bromide. In addition to this type of *in vivo* transposition, it is possible to use *in vitro* transposition, such as that mediated by the EZ::TN transposon tool (Epicentre, Madison, WI). This method eliminates the laborious *in vivo* transposition procedures, although the cost of the transposase is a limiting factor for large-scale experiments.

We examined transformation efficiency using shuttle mutagenesis. NHI-GT and NHI-ET were used for non-homologous integration to compare the efficiency of homologous integration with the trap elements HI-GT and HI-ET (Fig. 1c-f). DNA fragments containing no Physcomitrella sequences were integrated into

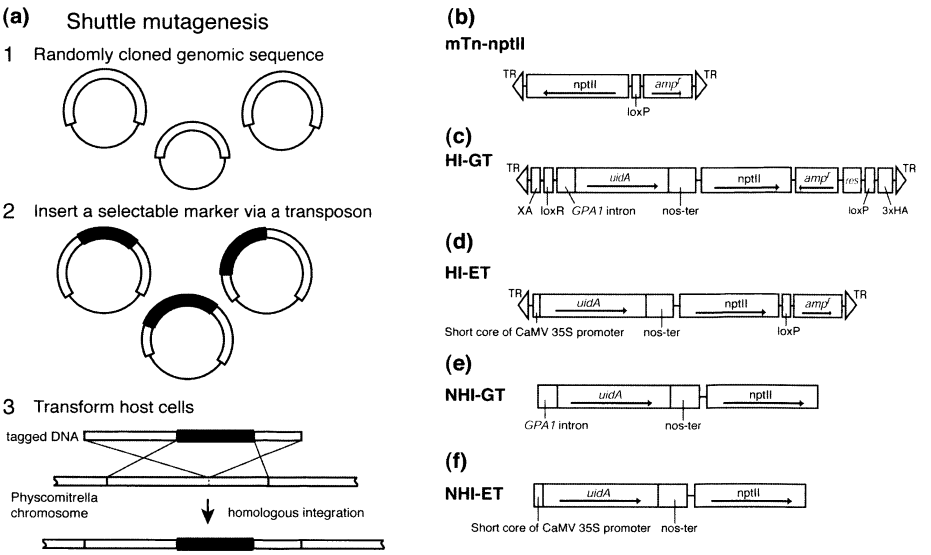

Figure 1. Schematic diagrams of the shuttle mutagenesis method and mini-transposons. The procedure for shuttle mutagenesis is based on three steps. (a) Physcomitrella genomic DNA sequences and mini-transposons are depicted as open and black boxes, respectively. Mini-transposon elements for gene tagging (b), gene-trap (c) and enhancer-trap (d) systems are shown. Cassettes for non-homologous integration of gene-trap (e) and enhancer trap (f) elements are shown. Gene-trap elements contain the Arabidopsis GPA1 intron. Enhancer-trap elements contain a portion of the CaMV 35S promoter. nptII indicates an expression cassette consisting of the CaMV 35S promoter, the nptII gene and a CaMV polyadenylation signal. mTn-nptII, HI-GT and HI-ET are mini-transposons containing 38-bp terminal repeats of the Tn3 transposon (TR), a loxP site, and an ampicillin resistance gene (amp^r). Arrows indicate the directions of transcription units.

the genome by non-homologous integration. Homologous integration with HI-GT and HI-ET generated approximately 10-fold more stable transformants than did non-homologous methods with NHI-GT and NHI-ET (Hiwatashi et al. 2001). We observed similar transformation efficiencies with other transposon constructs for homologous integration *versus* non-homologous integration (Nishiyama et al. 2000). This difference in transformation efficiency concurred with the results of a previous study Schaefer and Zrÿd 1997). These results suggest that high homologous integration frequency is based on recombination events that are dependent of homologous stretches in the transposon-inserted genomic DNA. Thus, although shuttle mutagenesis is a complicated method, it generates a larger number of tagged mutants than do methods mediated by non-homologous recombination.

More than 20,000 tagged lines have been generated in the past few years in our laboratory, and attempts to isolate the corresponding genes are in progress, although no gene has yet been isolated. Southern blot analysis to estimate the copy number of transposons inserted into each tagged line has shown that most lines contain multiple tags integrated into the genome. For example, 94% of 32 lines tagged with the HI-GT transposon were found to contain multiple insertions

(Hiwatashi et al. 2001). PCR, using primers that oriented in the distal direction of the transposon, amplifies DNA fragments of different sizes from each line, indicating that multiple copies of the transposons are integrated in close proximity (Nishiyama et al. 2000). We sequenced these PCR products from a tagged line, TN1, in which the gametophore is more stunted than in the wild type. We found that thirteen of the sixteen clones sequenced were derived from a 4.1-kb genomic clone, and three clones were derived from the cloning vector that was used for the genomic library construction. It is possible that a tandem repeat produced by extra-chromosomal homologous recombination among the introduced DNA fragments integrated at a single site (Nishiyama et al. 2000). Although this 4.1-kb genomic fragment likely contained a gene encoding a putative protein, disruption of this gene could not reproduce the stunted phenotype. It is possible that a gene in a region other than the 4.1-kb fragment was also disrupted by integration of another tag, and not all tag flanking sequences amplified by PCR may have been recovered. In addition, analysis of gene-trap lines suggests that tagged DNA fragments can integrate in non-homologous sites, as described in Section 2.3. If multiple copies of a transgene integrated into non-homologous sites, and this led to silencing of the expression of the endogenous gene (Flavell 1994), it would be impossible to identify the endogenous gene based on the sequence adjoining the tag. In the second line analyzed, various genomic fragments were isolated by "inter-transposon" PCR. Although multiple copies of transposons appeared to be integrated as a tandem repeat, different, seemingly unrelated, genomic DNA sequences were integrated along with the transposon and vector sequence. Thus, transformation by shuttle mutagenesis resulted in multiple tag integration events in tandem at a single genomic site or at multiple sites. This complicates the isolation of sequences flanking the tag by genomic sequence-based approaches. The integration of multiple copies of genomic sequences may also cause epigenetic effects, further complicating the identification of the gene that causes the phenotype. Reduction of the integrated transposons by Cre-*loxP* mediated recombination (Fig. 1b-d) may be a way to overcome this problem (Gu et al. 1993).

We have therefore taken another approach to generate a large number of tagged lines, containing fewer tags, using the restriction enzyme-mediated integration (REMI) technique. This technique was developed for *S. cerevisiae* and then applied to *Dictyostelium discoideum* and filamentous fungi, which are organisms with high rates of homologous recombination. As demonstrated in several organisms, REMI shows only a low frequency of multiple integration events of transforming DNA (Bolker et al. 1995, Granado et al. 1997, Kuspa amd Loomis 1992, Mullins and Kang 2001, Riggle and Kumamanto 1998, Schiestl RH, Petes 1991). REMI involves transforming cells with a linearized plasmid as a tag together with the restriction enzyme that was used for the linearization, and the tag is not accompanied by genomic sequences, which made it complicated to identify the disrupted gene in shuttle mutagenesis. It is assumed that the enzyme penetrates the cells and nuclear membranes and cuts chromosomal DNA at its specific restriction site. The linearized plasmid DNA, which has complementary ends, can be ligated to the chromosomal DNA ends by host cell enzymes, resulting in reconstitution of the restriction sites through the insertion of the plasmid. The integration of the plasmid tag into a particular gene depends on the presence of the particular restriction site in

the gene, so it is necessary to use several enzymes to generate a large number of insertion lines by REMI. The technique has been used to isolate a number of genes involved in development and pathogenicity (Loomis 1996, Riggle and Kumamanto 1998).

To investigate the potential usefulness of this technique in Physcomitrella, we analyzed the effect of REMI on PEG-mediated transformation of Physcomitrella protoplasts. A heterologous vector containing *nptII* as a selection marker was linearized with *Hin*dIII, *Eco*RI, or *Bam*HI and introduced into cells with or without different amounts of the same restriction enzyme. Our preliminary observations showed that addition of the restriction enzymes increased transformation frequencies, but the level of enhancement appeared to depend on the enzyme and the amount used (Y. Sato et al., unpublished results). *Bam*HI showed the largest increase, with up to 34-fold greater transformation frequency as compared to the transformants generated without adding enzyme. Only a three- to five-fold increase was observed with *Eco*RI and *Hin*dIII, respectively, although we believe further optimization of transformation by REMI is necessary. The inclusion of heat-denatured preparations of the same enzyme used to digest the plasmid did not increase the transformation rate. These results indicate that at least three restriction enzymes, in active form, increase the frequency of integration of linear DNA into Physcomitrella cells, encouraging further research into the potential of REMI in large-scale insertional mutagenesis. We are currently performing Southern blot analyses to analyze individual integration events, to determine whether the lines contain single or multiple integrated copies of the plasmid, and whether plasmids digested with a particular restriction enzyme are integrated into corresponding restriction sites, as in typical REMI integration events.

2.2. Gene- and Enhancer-Trap Lines

Gene- and enhancer-trap lines were generated using constructions that are variants of the transposon used in shuttle mutagenesis (Fig. 1c and d). Each element contains the nptII expression cassette and a reporter cassette for the *uidA* gene, which encodes β-glucuronidase (GUS) (Jefferson et al. 1987). The fourth intron of Arabidopsis *GPA1*, which encodes the alpha subunit of the G-protein gene, was fused to the 5' end of the reporter gene in the gene-trap elements (HI-GT) to enable detection of gene expression when the gene-trap element was inserted into either intron or exon sequences (Sundaresan et al. 1995). The enhancer-trap element (HI-ET) contains a short core sequence (–46 to +1) of the CaMV 35S promoter as a weak promoter (Benfey et al. 1989), fused to the 5' end of the *uidA* gene. We have also developed other elements with slight modifications, such as using a different spliced acceptor (Nishiyama et al. 2000).

Approximately 12,000 transformants containing gene-trap and enhancer-trap elements were screened for GUS activity, and 3-30% of stable transformants showed histochemical GUS staining. As shown in Fig. 2a-e, GUS activity was detected in various portions such as protonemata, gametophores, antheridia, archegonia, and sporophytes, indicating that genes or enhancers that function in many types of cells and tissues were trapped by this method. It is notable that axillary hairs were stained at much higher percentages than other cells or tissues

(Fig. 2e). Axillary hairs arise at the adaxial base of a leaf and are lined with several cells. These hairs are reported to produce mucilage in some mosses (Ligrone 1986), although this has not been determined in Physcomitrella. The higher percentages of GUS activity might reflect that genes expressed in the hairs or their promoters retain hot spots for the integration of the trap elements or, alternatively, that axillary hairs are cells in which proportionally more genes are expressed than in other cells and tissues. Also among the tagged lines were mutants with altered developmental and morphological phenotypes. Several lines showed altered responses to the phytohormones auxin or cytokinin, and therefore may have mutations in genes responsible for hormone perception or signal transduction (Fig. 2f and g).

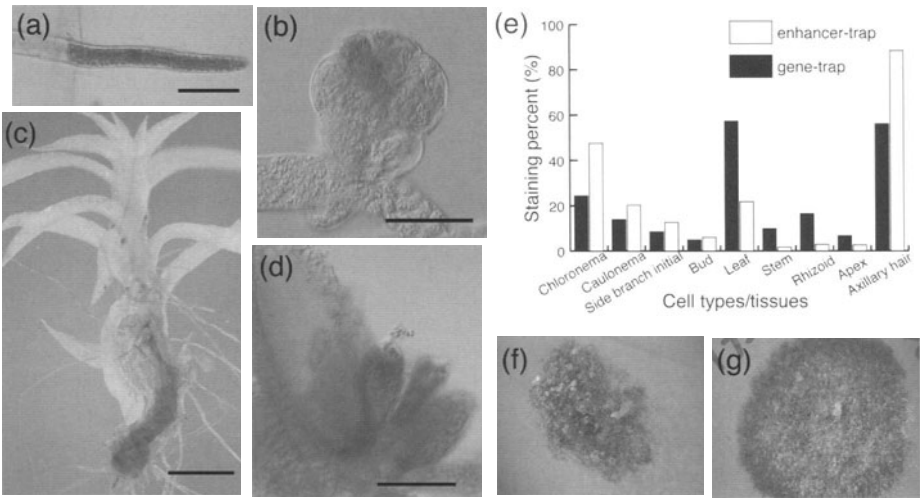

Figure 2. GUS expression patterns of gene-trap and enhancer-trap lines, and a tagged mutant with an altered hormone response.Staining of a rhizoid apical cell of gene-trap line YH87. (b) A bud of a young gametophore of gene-trap line YH229, showing staining of an apical cell and its surrounding cells . (c) A gametophore of gene-trap line YH209, showing staining of the basal region of the gametophore, where rhizoid filaments emarged. (d) A shoot apex of gene-trap line YH440, showing staining of an antheridium. Bars in (a) and (b) = 50 μm, in (c) = 500 μm, and in (d) = 100 μm. (e) Patterns of expression of the uidA gene in gene-trap and enhancer-trap lines generated with the homologous integration method. Y-axis values indicate the percentage of lines with expression in the indicated portions relative to the number of GUS-positive gametophytes, which was set at 100%. Apex indicates the number of lines displaying GUS staining in an apical cell and its surrounding cells. (f) A month-old wild-type gametophyte grown on medium supplemented with 1 mM BA, showing malformed bud formation. (g) A month-old gametophyte of tagged line 8617-7, which is resistant to cytokinin, grown on the same medium as the wild type.
(See colour section page x)

As multiple insertions of the transposons were observed in the trapped lines as well as in the tagged lines, isolation of a trapped gene, promoter or enhancer using a genomic sequence-based approach such as PCR may be possible but will probably be difficult. However, isolation of a trapped gene from a gene-trap line can be accomplished easily using the RACE (rapid amplification of cDNA ends)

method, as a transcript of a trapped gene are fused to that of the reporter gene. For an enhancer-trap line, the 5'-RACE method is amenable to isolation of a regulatory element and a trapped gene as long as a genomic sequence fused with the *uidA* reporter gene is transcribed. We describe examples of recovering trapped genes in the following section.

2.3. Isolation of Trapped Genes from Gene-Trap Lines

We have successfully isolated trapped genes from four individual gene- and enhancer-trap lines. However, an attempt to isolate trapped enhancers from enhancer trap lines was unsuccessful. Here we discuss the procedure for isolating trapped genes from gene-trap lines. In line YH209, which was generated using HI-GT, GUS activity was predominantly detected in mature rhizoid cells proximal to the stem (Fig. 2c) (Hiwatashi et al. 2001). Southern blot analysis of this line using a *uidA*-specific probe showed that YH209 contains at least four copies of the mini-transposon (Fig. 3a). Northern blot analysis of YH209 using the *uidA* probe revealed a broad band of around 2.4 kb (Fig. 3b), indicating that a portion of a Physcomitrella gene was fused to the *uidA* gene, as the detected band was longer than the expected 2.0-kb size of the *uidA* mRNA. Using *uidA* gene-specific primers and cDNA derived from YH209 gametophyte poly(A)$^+$ RNA, 5'-RACE was performed and cloned fragments were sequenced. The cloned fragments represented two sequences with splice variations. This suggests that the broad band detected in Northern blot analysis is probably caused by these splice variants. Subsequently, using one of the clones as a probe, an HI-GT-tagged DNA fragment termed 209-8SC was obtained from the transposon tagged genomic DNA library that was used to generate the Physcomitrella gene-trap lines. Sequencing analysis of the genomic fragment revealed that both representative sequences obtained by 5'-RACE were splice variants of the 209-8SC gene. Southern blot analysis of YH209 genomic DNA, using a portion of the 209-8SC gene as a probe, showed that two tagged 209-8SC fragments were integrated into different non-homologous sites. The 209-8SC fragment was re-introduced into wild-type Physcomitrella by homologous recombination. GUS staining of the lines in which a single copy of 209-8SC was integrated revealed expression patterns that were similar to those of the line YH209, except for additional activity at axillary hairs. Although the slight difference in the staining patterns may be due to the lack of a complete regulatory element in the 209-8SC fragment, the similarity of the expression patterns indicates that the GUS activity of line YH209 reflects the endogenous expression pattern of the gene encoded on 209-8SC.

Finally, 3'-RACE performed using the sequence of the clones obtained through 5'- RACE revealed the sequence of the gene located on 209-8SC. The gene was found to encode a putative protein of 916 amino acids with strong similarities to plant acidic alpha-glucosidase (Hiwatashi et al. 2001). Proteins in the plant alpha-glucosidase family are thought to function in starch degradation or xyloglucan oligosaccharide metabolism. Since the rhizoid is probably a non-photosynthetic organ, as they do not contain developed chloroplasts, the identified protein is probably involved in xyloglucan oligosaccharide metabolism. Sequencing of 209-

8SC showed that the fragment also contains a regulatory region that is responsible for the expression pattern of the acidic alpha-glucosidase gene.

Using the same approach, we isolated three more trapped genes from gene-trap and enhancer-trap lines, in which these genes were expressed predominantly in the apical regions of protonemata and gametophores. Southern blot analysis of line YH229 (Fig. 2b), which was generated with an HI-GT transposon, revealed integration of multiple transposon copies (Fig. 3c). However, Northern blot analysis with a GUS reporter gene as a probe detected three transcripts longer than a GUS transcript were in this line (Fig. 3d). Subsequently, 5'-RACE was performed, and one of the resulting fragments was used to identify the trapped fragment in the original transposon-tagged genomic DNA library. The identified trapped fragment was re-introduced into the wild type by homologous recombination, and the resulting transformants had expression patterns similar to that of the original gene-trap line. The fragment was found to contain a promoter region and a gene encoding a member of the plant-specific kinesin-like protein family. We found that 5'-RACE can efficiently identify trapped genes in gene-trap lines, even when multiple tags are present in the Physcomitrella genome. Southern blot analysis showed that in the original YH229 line, as in the YH209 line, non-homologous integration occurred. The integration into non-homologous sites might have been caused by extrachromosomal homologous recombination among a mixture of transposon-tagged genomic DNA fragments that we used in the transformation during the generation of trap lines (Hiwatashi et al. 2001).

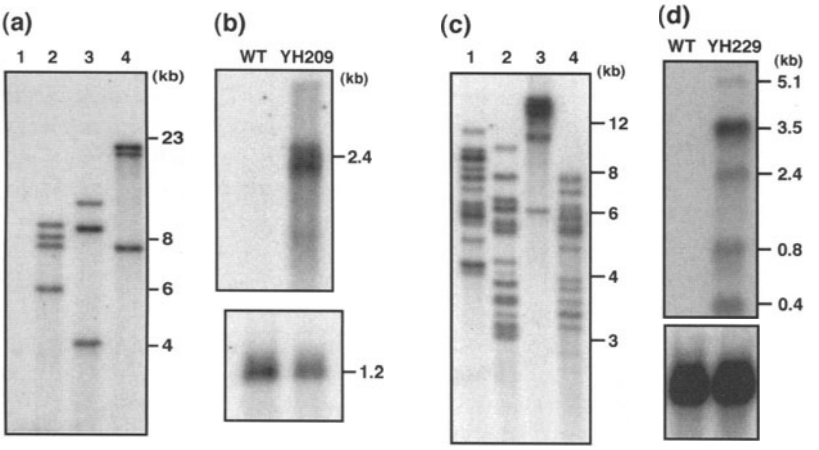

Figure 3. Southern and Northern blot analyses of trapped lines *(a) Southern blot analysis of wild type (lane 1) and line YH209 (lanes 2 to 4). Blots were probed with a uidA gene-specific probe. Genomic DNA was digested with EcoRI (lanes 1 and 3), BglII (lane 2) and HindIII (lane 4). (b) Northern blot analysis of wild type and line YH209, using a uidA probe. Southern blot (c) and Northern blot (d) analyses of line YH229, using a probe for the uidA gene. Genomic DNA was digested with BglII (lane 1), EcoRI (lane 2), HindIII (lane 3), or PstI (lane 4). The lower panels in (b) and (d) show blots probed with a glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene as a control.*

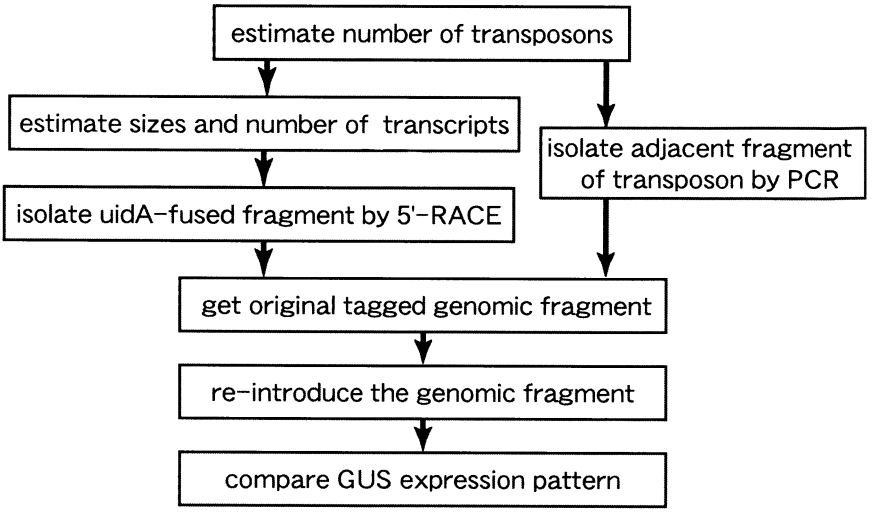

Figure 4. Experimental outline for the identification of a trapped gene.

The procedure for identification of a trapped gene from gene- and enhancer-trap lines, summarized in Fig. 4, is as follows. (1) Estimate the number of integrated transposons in the trapped line using Southern blot analysis. If only one or a few transposons are detected, isolate the genomic regions that border the tag using a direct PCR-based method, or perform the following steps. (2) Estimate the sizes and the number of the candidate transcripts containing both the trapped gene and the GUS reporter gene, using Northern blot analysis. (3) Isolate the *uidA*-fused cDNA fragments from the trap line with 5'-RACE using *uidA*-specific primers. (4) Isolate the corresponding transposon-tagged DNA fragment from the original tagged genomic DNA library. (5) Re-introduce the isolated tagged DNA fragment into the wild type by homologous recombination. (6) Compare the GUS expression patterns of the obtained transformants with those of the original trap line.

As examined in the YH209 and YH229 lines, tagged genomic DNA fragments were integrated at non-homologous sites. Integration in a non-homologous site keeps the originally targeted site from being disrupted by a trapped element. This is probably advantageous for generating trapped lines in a moss system. Because Physcomitrella is in the autotrophic haploid generation for the majority of its life cycle, the disruption of a gene or a regulatory element essential for cell viability might be lethal. Integration at a non-homologous site would reduce the possibility of such deleterious effects. If a regulatory element that helps control expression of the reporter gene is present in the tagged genomic DNA fragment, it is not necessary that gene-trap or enhancer-trap elements integrate at the original site. Integration into non-homologous sites tends to produce mutants containing multiple and various insertions, which would increase the possibility of trapping genes and regulatory elements.

The suitability of Physcomitrella in the study of cellular and developmental processes has long been recognized, as this moss has a simple structure with a small

number of cell types, enabling observation of individual cells (Cove et al. 1997, Reski 1998, Schumaker and Dietrich 1998). The establishment of gene-trap and enhancer-trap systems enables the identification of genes or regulatory elements based on their spatio-temporal cellular expression patterns.

3. GENERATION AND QUALITY OF FULL-LENGTH cDNA RESOURCES

3.1. Construction of Full-Length Physcomitrella cDNA Libraries

The development of a wide range of genetic tools for Physcomitrella is essential for functional analyses of genes. We generated high-quality full-length Physcomitrella cDNA libraries. In the following sections we describe the construction and quality of these libraries, the analysis of expressed sequence tags (ESTs) from the libraries, and an example of the utilization of the libraries for a functional screen.

For the development of the full-length enriched Physcomitrella cDNA libraries, we used recent technical advances that allowed us to preferentially clone cDNA that corresponds to a full-length mRNA with a cap structure at the 5' end (Kristiansen and Pandey 2002). Reverse transcription reactions were carried out at a high temperature, using trehalose as a stabilizer, to disrupt the secondary structure of the RNA and favor the completion of the reverse transcription reaction. Full-length cDNA-RNA hybrids were then recovered on an avidin column through a biotinylated cap structure on the molecule. The same method has been used previously to create high-quality full-length enriched mouse and Arabidopsis cDNA libraries (Kawai et al. 2001, Selki et al. 2002). Similar methods are also available for preparing full-length cDNA libraries, and these methods have been successfully used in large-scale analyses of human (Suzuki et al. 2001), *Drosophila melanogaster* (Stapleton et al. 2002) and rice (Osato et al. 2002) sequences. Full-length cDNAs are essential not only for gene annotation and the determination of transcriptional start sites, but also for functional analyses of genes and their products, such as protein-protein interactions as studied in mouse (Suzuki et al. 2001). The replacement of an original gene with various mutated versions will help to define the relationship between the structure of a gene and its function, using homologous recombination, which is inefficient in other plants but can be used as a powerful tool in bacteria and yeast (Schaefer 2002).

3.2. Physcomitrella EST Database

We constructed three full-length enriched cDNA libraries from Physcomitrella gametophytes that were either not treated or treated with naphthalene acetic acid (NAA; auxin) or benzyladenine (BA; cytokinin) (Nishiyama et al. 2003). The latter two libraries were normalized using a subtraction procedure to reduce the representation of highly expressed mRNAs (Carninci et al 1996). We isolated 14,348, 19,200, and 18,432 clones, respectively, from these libraries, and sequenced their inserts from both ends, obtaining 85,191 total EST sequences. The collection of EST data is highly important in organisms whose genomes have not yet been sequenced, as it allows the rapid identification of expressed genes. The first Physcomitrella EST projects were reported in the late 1990's by Reski et al. (1998)

and Machuka et al. (1999). These efforts have been continued by larger-scale Physcomitrella EST analyses by the Physcomitrella EST programme (PEP) [http://www.moss.leeds.ac.uk/], Reski's group (Rensing et al. 2002), and our project, PHYSCObase [http://moss.nibb.ac.jp/]. More than 100,000 ESTs should be publicly available by the end of 2003, and will represent one of the largest plant EST collections available. Clones corresponding to individual EST sequences are available from each group. Our full-length clones are available from the RIKEN Bio Resource Center [http://www.brc.riken.go.jp/lab/epd/Eng/].

Compiling our EST data with the Physcomitrella mRNA sequences already available in public databases created a collection of 102,553 sequences. We assembled these sequences into 19,775 contigs from our libraries and 2,290 contigs from other sources. Since each clone was sequenced from both ends, the 5'- and 3'-contigs could both be assigned to the original clone, a process that condensed the 19,775 contigs into 13,593 pairs of contigs. We assigned to each pair of contigs a pair ID (i.e., Pnnnnnn, where n is a digit). The pairs were designated as putative transcripts, since each pair probably corresponds to an independent mRNA. If one considers the 2,290 contigs from the other sources as independent transcripts, a total of 15,883 putative transcripts can be catalogued as Physcomitrella transcripts. The data are summarized in Table 1 and can be browsed through PHYSCObase. All sequences and contig data, as well as a BLAST search program for these sequences, are available at this database. We designated clones derived from non-treated, NAA-treated and BA-treated libraries as pph, pphn and pphb, respectively. Consensus sequences based on each contig, and the results of BLASTX comparisons of clones, contigs and pairs with a non-redundant dataset, are available in the database.

Table 1. Full-length Physcomitrella clones from three cDNA libraries

Treatment	Harvest (days)[a]	Tissues	Clone Name	5' EST	3' EST	Pairs
None	13-14	chloronema-rich	pph	9,937	10,309	5,955
Auxin	8-11	caulonema-rich	pphn	16,733	16,761	6,601
Cytokinin	8-13	malformed bud-rich	pphb	16,450	15,001	7,851
Total				43,120	42,071	13,593

[a] *Tissues were cultured for the indicated periods on medium supplemented with 1 μM NAA (auxin) or 0.5 μM BA (cytokinin), or without hormones (none).*

3.3. Quality of the Libraries

The EST data were utilized in two ways to evaluate the quality of the full-length cDNA libraries: by analyzing both the proportion of clones that contained a complete protein coding sequence, and the proportion of full-length cDNA clones that contained a transcriptional start site. For the former criterion, 31 genes with complete protein coding sequences were selected from GenBank and compared with corresponding EST clones from our libraries to observe whether the clones had complete 5' and 3' ends (Nishiyama et al. 2003). This revealed that 95.4% of the clones contained 5' upstream regions, including either a start codon or a 5' untranslated region, and 97.8% of clones contained 3' downstream regions, including either a stop codon or a 3' untranslated region. Assuming that 5'-end and 3'-end truncation occurs independently, the proportion of clones that have both start and stop codons can be calculated as the product of the proportion of clones that are complete at their 5'- and 3'-ends. Using this formula, we estimated that 93% (95.4% × 97.8%) of the clones contain complete coding sequences.

In order to estimate the proportion of full-length cDNA clones that contain a transcriptional start site, we utilized the fact that transcription usually starts at a purine base (Lewin 2000). Among our ESTs, the base composition of the first nucleotide was strongly biased for purine, at 81%, whereas 42-44% of the second to twentieth nucleotides were purine (Fig. 5). If we assume that all the pyrimidine bases (19%) present as the first nucleotide in our ESTs were caused by a truncation of the mRNA, 14% of the first nucleotides are calculated to be purine derived from the truncation of a transcriptional start site (14% = 19% × 42/58; 42/58 = purine to pyrimidine ratio at the second through twentieth positions). This leads, in turn, to the estimation that 67% (100-19-14) of the first nucleotides represent a transcription start site. A low proportion of pyrimidine bases at transcriptional start sites has been reported (Bannerjee 1980), and a certain proportion of the observed pyrimidine bases is likely to be present under the first base at the cap structure. We must caution, however, that this estimation is only speculative, since it has been reported that reverse transcriptase can add a few cytosine nucleotides at the 3' end of the first strand of the cDNA (Schmidt and Mueller 1999). A proportion of guanine nucleotides must have been counted as 5' guanine nucleotides, regardless of whether they actually represent the 5'-end of the cDNA. Nevertheless, the percentage of adenine nucleotides at the first position is clearly higher than at the second through twentieth positions, which again supports the idea that a significant proportion of the clones contain a transcriptional start site. Full-length cDNA libraries from Arabidopsis and mouse, which were made using essentially the same method as our Physcomitrella libraries, have been evaluated to contain 53% to 95% of full-length cDNA clones, based on comparison to the putative full-length sequence in databases Carninci et al 1996, Kawai et al. 2001, Seki et al. 1998).

Since we generated the three types of full-length libraries from gametophytes that were either untreated or treated with auxin or cytokinin, we analyzed the libraries for over-representation of auxin- and cytokinin-related transcripts (Nishiyama et al. 2003). These transcripts include genes that are regulated by auxin and cytokinin, as well as genes that are related to the

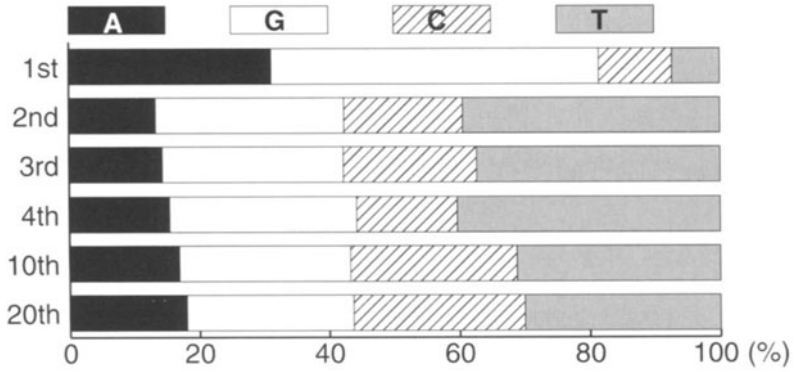

Figure 5. Base composition of each position in the 5' region of the cDNAs.

The proportion of each base, A, G, C and T was calculated at positions 1, 2, 3, 4, 10 and 20, and depicted as black, white, hatched, and gray boxes, respectively.

differentiation and development induced by auxin or cytokinin treatment. Sixty of these transcripts were over-represented in the auxin-treated library, as compared to the non-treated and cytokinin-treated libraries, at the 0.1% significance level. The over-represented transcripts in the auxin-treated library included many structural and stress-related genes, and appeared to be consistent with a set of genes reported to be induced by auxin in Arabidopsis (Qing et al. 2002). We also found 68 putative transcripts that were over-represented in the cytokinin-treated library, at the 0.1% level. These included transcripts that encode a member of the cytokinin oxidase family, receptor-like kinases, and EREBP-like transcription factors.

3.4. Resources to Elucidate Functions of Genes Common to Land Plants

We analyzed the proportion of Arabidopsis genes with counterparts that are expressed in moss gametophytes (Nishiyama et al. 2000). All Arabidopsis protein sequences were compared with the *Physcomitrella* dataset. As shown in Table 2, 66.4% (17,382/26,178) of the Arabidopsis genes have homologs in Physcomitrella, with an *E* value threshold of $1x10^{-3}$. Similarity searches were also performed against datasets from *D. melanogaster*, *S. cerevisiae*, and *Synechocystis* sp. PCC 6803, and putative homologs were discovered at frequencies of 47.1, 40.5, and 21.3%, respectively ($E \leq 1x10^{-3}$). This suggests that the Arabidopsis genome contains more genes homologous to Physcomitrella genes than to gene sets from other organisms such as protozoa, fungi, and metazoa.

Ten of sixteen classes of plant-specific transcription factors inferred from the Arabidopsis genome (AGI 2000), were identified in Physcomitrella: the MYB R2R3, Myb single-repeat protein, homeodomain-leucine zipper, ARF-AUX/IAA, NAC-like, zinc finger-DOF, zinc finger-WRKY, AP2/EREBP, ABI3/VP1-like, and GRAS-like transcription factors. The remaining six classes (cycloidea-like, YABBY, EIN3, RAV-like, squamosa-binding protein, and GT1-like) have not yet been identified in the Physcomitrella EST dataset. Certain widespread transcription factors (REL-homology region protein, nuclear steroid receptors, forkhead-winged

Table 2. Comparison of the Arabidopsis protein dataset to protein datasets of various organisms

Organism	Ratio (%)
P. patens	66.4
D. melanogaster	47.1
S. cerevisiae	40.5
Synechocystis sp. PCC 6803	21.3

[a] *Similarity searches were performed with TBLASTN using all Arabidopsis protein sequences in the NCBI RefSeq database as the query. The E value of the best hit from each search was recorded. The number of clones that showed significant similarities ($E \leq 1x10^{-3}$) was divided by the total number of queries (26,178), and the results are shown as percentages.*

helix, and POU domain families) and key signal transduction regulators (i.e., TGF-β/SMADs, JAK/STAT, Notch, and Wingless/Wnt) in metazoa were found in neither the Arabidopsis genome (AGI 2000) nor in the Physcomitrella dataset. Since the Physcomitrella ESTs probably do not represent all of the transcripts produced from the Physcomitrella genome, the number of Arabidopsis genes homologous to Physcomitrella genes should be higher than the current estimate of 17,382. Because the functions of many predicted Arabidopsis genes and proteins are still unknown, functional analyses of their homologs in Physcomitrella using gene targeting techniques should accelerate our understanding of the functions of genes common to land plants.

3.5. ESTs Present in Physcomitrella but not in Vascular Plants

Although the Physcomitrella transcriptome appears similar to that of Arabidopsis, we have also identified a set of genes that appear to be present in Physcomitrella but not in vascular plants (Nishiyama et al. 20003). A BLASTX search of putative Physcomitrella transcripts against the non-redundant database of amino acid sequences deposited in GenBank revealed that 10,757 out of 15,883 had significant similarities to GenBank sequences ($E \leq 1x10^{-3}$). Among these, 9,907 transcripts were highly similar to genes of vascular plants, whereas 850 transcripts had high similarity to genes of other organisms. As some of the 850 transcripts showed similarities to genes from both vascular plants and other organisms, with

low levels of difference, the significance of which could not be defined, we selected 300 of the 850 putative transcripts with *E* values that were at least $1x10^{7}$-fold lower than the *E* values for the highest similarity to genes of vascular plants. We termed these 300 putative transcripts MSTs (moss transcripts absent in vascular plants). The 300 MSTs are candidates for transcripts that are absent in vascular plants but present in mosses, and probably correspond to previously undiscovered land plant genes. Some of the 300 MSTs may represent genes that were lost or diverged rapidly in vascular plant lineages, or that were incorporated into Physcomitrella by horizontal gene transfer from other organisms. Thus, the genes corresponding to the MSTs have probably either acquired new functions or are not required in vascular plants, and are likely to be involved in moss-specific processes. Some of the MSTs have similarities to genes that encode transporters, two-component histidine kinases, or DNA damage repair enzymes predominantly found in bacteria and fungi. Analysis of the DNA damage repair enzyme genes may help reveal the mechanisms that govern the highly efficient homologous recombination of Physcomitrella.

Examination of ESTs that are highly represented in the non-treated library, which may be genes with high expression levels in Physcomitrella gametophytes, revealed that seven of the thirty most highly expressed genes are related to photosynthetic pathways. This indicates that the tissues used to prepare the library are active in photosynthesis. In addition, two MSTs with similarity to a high-affinity iron permease from *Candida albicans* (E=$3x10^{-48}$) and a putative phosphate permease from *Chlamydomonas reinhardtii* (E=$5x10^{-43}$) were among the top thirty. These two highly expressed MSTs, which probably have roles in ion transport, may reflect the importance of this process in moss gametophytes, as compared to vascular plants.

The BLASTX search revealed that about 30% of the Physcomitrella transcripts exhibited no significant similarity to genes from other organisms (although only the end sequences were determined for these). It is likely that some of these transcripts represent additional MSTs. MSTs will be promising targets in the search for plant gene products with new functions or activities that offer possible future applications in agriculture or medicine, in addition to the contributions they can make to basic science.

3.6. Redundant Gene Sets in Physcomitrella

Genome-wide duplication and local gene duplication have been reported to have occurred in a large proportion of the Arabidopsis genome (AGI 2000). Molecular biological studies in Physcomitrella have demonstrated that many Physcomitrella genes are closely related to other genes in the same organism. To assess the level of gene redundancy in Physcomitrella, we performed a BLASTN search to find the most closely related contig to each contig in the Physcomitrella dataset. The identity of the second best hit to each query contig was recorded, since the best hit is the gene originally queried. About 28% of the contigs contained sequences with strong similarity (identity $\geq$ 80%) to other contigs in a stretch greater than 100 bp. The same procedure was performed on the Arabidopsis mRNA dataset, and 38% of the gene set was highly related to other Arabidopsis genes. These results indicate that both organisms have a partly redundant genome.

We compared the synonymous evolutionary distances (Ks) between seven published closely related Physcomitrella gene pairs (Table 3). The values ranged from 0.0006 to 0.23, a large range that suggests that gene duplication occurred at different dates. The gene duplications of *PpRAD51A* and *PpRAD51B,* and *PpLFY1 and PpLFY2,* may have occurred at a similar time. However, the gene duplication of *PpCry1a* and *PpCry1b* probably occurred much more recently than did the duplication of *PpLFY1* and *PpLFY2*, whereas the divergence event of *PpRPOT1*and

Table 3. Synonymous substitutions between Physcomitrella sister genes

Names of sister genes (gi accession number)		Ks[a]	S.D.[b]
PpLFY1 (AB052251.1)	*PpLFY2* (AB052252.1)	0.0208 6	0.0050 8
PpRAD51A (AJ316537.1)	*PpRAD51B* (AJ316538.1)	0.0297 9	0.0062 4
PpRPOT1 (AB055214.1)	*PpRPOT2* (AB055215.1)	0.2365 5	0.0120 2
MKN2 (AF285147.1)	*MKN4* (AF284817.1)	0.1696 1	0.0214 8
PpCry1a (AB027528.2)	*PpCry1b* (AB060693.1)	0.0006 0	0.0006 0
PpMADS2 (AB067689.1)	*PpMADS3* (AB067690.1)	0.0672 7	0.0110 7
PPM1 (AF150931.1)	*PPM2* (AF150933.1)	0.1284 8	0.0151 9

[a] *Synonymous evolutionary distance (Ks) was calculated using LWL91 (Liu et al. 1985).*
[b] *Standard deviation.*

PpRPOT2 is probably a relatively ancient event. If the closely related genes were created at the same time by genome duplication, the Ks values would be expected to be much more similar. These data suggest that many of the gene duplication events occurred independently, which led to the current redundancy of the Physcomitrella genome. Although we cannot rule out the possibility of an ancient genome duplication or polyploidization, the redundancy of the present Physcomitrella genome cannot be entirely explained by genome-wide duplication. The sequencing of the Physcomitrella genome should shed light on the question of genome duplication.

4. FULL-LENGTH cDNA OVEREXPRESSION AS A TOOL FOR GAIN-OF-FUNCTION SCREENING

4.1. Systematic Overexpression Screening to Discover New Genes

We used gain-of-function screening to complement loss-of-function screening for the discovery of new gene functions using large-scale overexpression screening with Physcomitrella full-length cDNAs. Figure 6a shows the procedure. Generation of high quality full-length clones and advances in mass-sequencing techniques were crucial for this screening. Each full-length cDNA selected for overexpression was subcloned under the control of a constitutive promoter (Fig. 6b), and the resultant plasmid was transformed into Physcomitrella protoplasts to allow transient overexpression of the genes during regeneration. The vector contains a GFP expression cassette (Chiu et al. 1996), which allows transformed cells to be distinguished from untransformed cells, based on green fluorescence. We screened for abnormal phenotypes during protoplast regeneration. The protoplast regeneration involves many points of biological interest, such as cell polarity, asymmetric cell division, and cell differentiation (Ashton et al. 1990, Cove 2000). We are able to observe regeneration up to the 20-cell stage of regenerated protonema, since GFP fluorescence is transiently maintained until this stage. This screening system has the advantage that any abnormal phenotypes observed can be immediately attributed to the gene used in the transformation, allowing causal genes to be rapidly identified. The identified cDNAs can be selected for functional analysis, which is straightforward in Physcomitrella due to the ease of gene targeting analyses (Schaefer 2002).

This overexpression screening through individual genes not only complements loss-of-function screening, such as the aforementioned gene-tagging systems, but also has several advantages over other gain-of-function methods, such as activation tagging. Activation tagging, which uses T-DNA vectors that harbor CaMV 35S enhancers to "activate" genes, has been successfully used to identify developmental genes in model angiosperms (Jeong et al. 2002, Walden 2002, Weigel et al. 2000) However, using this system, many genes known to result in a defined phenotype when expressed under the control of a constitutive promoter were not identified, even when a number of enhancer-tagged mutants potentially representing an almost complete saturation of the genome, were screened (Weigel et al. 2000). Results such as these can perhaps be explained by the presence of hot spots into which the tagged sequence integrates preferentially, or insulator sequences, which block the effects of

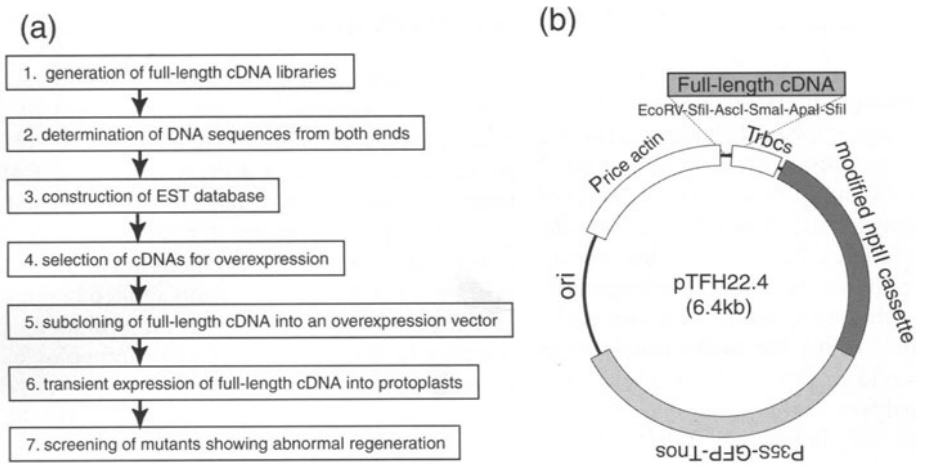

Figure 6. (a) Strategy of gain-of-function screening by overexpression of full-length cDNAs. (b) The pTFH22.4 overexpression vector. Putative full-length clones are inserted into appropriate multicloning sites under the control of a constitutive promoter derived from rice actin. Included are an nptII cassette modified to function in both bacteria and plants and a GFP cassette to signal the incorporation of the plasmid into transformed protoplasts.

the enhancer on a nearby gene, resulting in apparent selectivity of enhancer activity. In our overexpression screening system, these problems are circumvented due to the constitutive overexpression of individual cDNAs, independently of the position of the genes in the chromosome and the sequences adjacent to the integrated gene. However, there are some limitations to our overexpression system. Even though we generated high-quality libraries of full-length cDNAs, about 10% of the cDNAs do not have a complete protein coding region, and this proportion is probably larger among longer cDNAs. In addition, cDNAs that represent genes with low expression levels are probably under-represented in the libraries. Complementary approaches should be applied in order to best understand a genome-wide network of genes.

Overall, the overexpression system worked well. To test the overexpression system, we transiently overexpressed cDNAs that encode actin and alpha-tubulin in isolated protoplasts. Both proteins are involved in cytoskeletal organization and regulate tip growth (Marthur 2002, Ringli 2002). Transient overexpression of actin resulted in abnormal regeneration of the protoplasts, manifested by clusters of clumped cells. Transient expression of alpha-tubulin resulted in clumped cells with bulbous apical cells, which is consistent with the bulged phenotype of apical cells treated with agents that target microtubules (Doonan 1988). These results suggest that this system will be effective for identifying genes responsible for protoplast regeneration. In further screenings, we found more than 100 cDNAs that affect regeneration of protoplasts. These genes encode proteins involved in cell cycle regulation, cell polarity, cell differentiation, cell death, and organelle development, and proteins of unknown function.

5. CONCLUDING REMARKS

In this chapter, we described the construction of high-quality full-length Physcomitrella cDNA libraries, as well as systems to identify genes using transposon-based tagged mutant lines, gene- and enhancer-trap systems, and overexpression screening. As the Physcomitrella genome is generally similar to that of Arabidopsis, and the functions of many Arabidopsis genes have yet to be defined, analyses of related Physcomitrella genes will provide routes for unraveling their functions. Novel genes that appear to be specific to Physcomitrella may also lead to the discovery of novel functions that will be useful in both basic and applied biology. Furthermore, because mosses and vascular plants diverged early in the evolution of land plants, the differences and similarities in their life cycles and morphologies should be represented in their genomes (Kenrick and Crane 1997), and comparative analyses of the genes of non-vascular and vascular plants should shed light on the diversity in land plants.

6. ACKNOWLEDGEMENTS

We would like to thank previous and present coworkers, our colleagues, and the technicians in our laboratory for their support. Research and preparation of this chapter were supported by a Grant-in-Aid for Scientific Research from MEXT and JSPS (to TF and MH). The rice actin promoter for overexpression was kindly donated by Dr. R. Wu (Cornell Univ., NY).

7. REFERENCES

Ashton. N.W., Boyd, P.J., Cove, D.J., & Knight, C. D. (1990) Protoplasts as tools in the study of moss development. In: Chopra, R.N.B. (Ed.) *Bryophyte development: physiology and biochemistry*, (pp. 1-16) CRC Press, Boca Raton, FL.

Bannerjee, A.K. (1980) 5'-terminal cap structure in eukaryotic mRNAs. *Micro. Rev.* 44, 175-205 (1980).

Bellen, H.J. (1999) Ten years of enhance detection: lessons from the fly. *Plant Cell* 11: 2271-2281

Benfey, P.N., Ren, L. & Chua, N.-H. (1989) The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. *EMBO J.* 8, 2195-2202

Bolker, M., Bohnert, H.U., Braun, K.H., Gorl, J., & Kahmann R (1995) Tagging pathogenicity genes in ustilago-maydis by restriction enzyme-mediated integration (REMI). *Mol. Gen. Genet.* 248, 547-552

Carninci, P., Kvam, C., Kitamura, A., Ohsumi, T., Okazaki, Y., Itoh, M., Kamiya, M., Shibata, K., Sasaki, N., Izawa, M., Muramatsu, M., Hayashizaki, Y., & Schneider, C. (1996) High-efficiency full-length cDNA cloning by biotinylated CAP trapper. *Genomics* 37, 327-336

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., & Hayashizaki, Y (2000) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10, 1617-1630

Chiu, W.L., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H,, & Sheen J. (1996) Engineered GFP as a vital reporter in plants. *Curr. Biol.* 6, 325-330

Cove, D.J. (2000) The generation and modification of cell polarity. *J. Expt. Bot.* 51, 831-838

Cove, D.J., & Knight, C.D. (1993) The moss Physcomitrella patens, a model system with potential for the study of plant reproduction. Plant Cell 5, 1483-1488

Cove, D.J., Knight, C.D., & Lamparter T. (1997) Mosses as model systems. *Trends Plant Sci.* 2, 99-105

Doonan, J.H., Cove, D.J., & Lloyd, C.W. (1988) Microtubules and microfilaments in tip growth: evidence that microtubules impose polarity on protonemal growth in *Physcomitrella patens*. *J. Cell Sci.* 89, 533-540

Egener, T., Granado, J., Cuitton, M.-C., Hohe, A., Holtorf, H., Lucht J.M., Rensing, S.A., Schlink, K., Schulte, J., Schween G., Zimmermann, S., Duwenig, E., Rak, B., & Reski, R. (2002) High frequency of phenotypic deviations in *Physcomitrella patens* plants transformed with a gene-disruption library. *BMC Plant Biol.* 2, 1-9

Flavell, R.B. (1994) Inactivation of gene-expression in plants as a consequence of specific sequence duplication. Proceed. Natl. Acad. Sci. U.S.A. 91, 3490-3496

Granado, J.D., Kertesz Chaloupkova, K., Aebi, M., and Kues U (1997) Restriction enzyme-mediated DNA integration in *Coprinus cinereus*. *Mol. Gen. Genet.* 256, 28-36

Gu, H., Zou, Y.R., & Rajewsky K (1993) Independent control of immunoglobulin switch recombination at individual switch regions evidenced through cre-*loxP*-mediated gene targeting. *Cell* 73, 1155-1164

Hiwatashi, Y., Nishiyama, T., Fujita, T., & Hasebe M. (2001) Establishment of gene-trap and enhancer-trap systems in the moss *Physcomitrella patens*. *Plant J.* 28, 105-116

Jefferson, R.A., Kavanagh, T.A., & Bevan, M.W. (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene marker in higher plants. *EMBO J.* 6, 3901-3907

Jeong, D.H., Kang, H.G., Moon, S., Han, J.J., Park, S., Lee, H.S., An, K.S., & An GH (2002) T-DNA insertional mutagenesis for activation tagging in rice. *Plant Physiol.* 130, 1636-1644

Kawai, J.S.A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M,, Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S., & Hayashizaki, Y. (2001) Functional annotation of a full-length mouse cDNA collection. *Nature* 409, 685-690

Kenrick, P., & Crane PR (1997) The origin and early evolution of plants on land. *Nature* 389, 33-39

Kristiansen, T.Z., & Pandey, A. (2002) Resources for full-length cDNAs. *Trends Biochem. Sci.* 27, 266-267

Kuspa, A., & Loomis, W.F. (1992) Tagging developmental genes in dictyostelium by restriction enzyme-mediated integration of plasmid DNA. *Proc. Natl. Acad. Sci. U.S.A.* 89, 8803-8807

Lewin, B. (2000) Genes VII. Oxford University Press, Oxford

Li, W., Wu, C., & Luo, C (1985) A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol Biol Evol* 2, 150-174

Ligrone, R. (19896) Structure, development and cytochemistry of mucilage-secreting hairs in the moss Timmiella barbuloides (Brid.) Moenk. *Annals Bot.* 58, 559-568

Loomis, W.F. (1996) Genetic networks that regulate development in Dictyostelium cells. *Microbiol Rev.* 60, 135-50

Machuka, J., Bashiardes, S., Ruben, E., Spooner, K., Cuming, A., Knight, C., & Cove D (1999) Sequence analysis of expressed sequence tags from an ABA-treated cDNA library identifies stress response genes in the moss *Physcomitrella patens*. *Plant Cell Physiol.* 40, 378-387

Mathur, J., & Hulskamp, M. (2002) Microtubules and microfilaments in cell morphogenesis in higher plants. *Curr. Biol.* 12, R669-R676

Mullins, E.D., & Kang, S. (2001) Transformation: a tool for studying fungal pathogens of plants. *Cell. Mol. Life Sci.* 58, 2043-2052

Nishiyama, T., Fujita, T., Shin-I, T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y., & Hasebe, M. (2003) Comparative genomics of the Physcomitrella gametophytic transcriptome and Arabidopsis genome: implication for the land plant evolution. *Proc Natl Acad Sci U S A: in press*

Nishiyama, T., Hiwatashi, Y., Sakakibara, K., Kato, M., & Hasebe, M.(2000) Tagged mutagenesis and gene-trap in the moss, *Physcomitrella patens* by Shuttle mutagenesis. *DNA Res.* 7, 9-17

Osato, N., Itoh, M., Konno, H., Kondo, S., Shibata, K., Carninci, P., Shiraki, T., Shinagawa, A., Arakawa, T., Kikuchi, S., Sato, K., Kawai, J., & Hayashizaki, Y. (2002) A computer-based method of selecting clones for a full-Length cDNA project: simultaneous collection of negligibly redundant and variant cDNAs. *Genome Res.* 12, 1127-1134

Qing, T., Nicholas, J.U., & Jason, W.R. (2002) Arabidopsis SHY2/IAA3 inhibits auxin-regulated gene expression. *Plant Cell* 14, 301-319

Rensing, S.A., Rombauts, S., Van de Peer, Y., & Reski, R. (2002) Moss transcriptome and beyond. *Trends Plant Sci.* 7, 535-538

Reski, R. (1998) Development, genetics and molecular biology of mosses. *Botanica Acta* 111, 1-15

Reski, R., Reynolds, S., Wehe, M., Kleber-Janke, T., & Kruse, S. (1998) Moss (*Physcomitrella patens*) expressed sequence tags include several sequences which are novel for plants. *Botanica Acta* 111, 143-149

Riggle, P.J., & Kumamoto, C.A. (1998) Genetic analysis in fungi using restriction-enzyme-mediated integration. *Curr. Opin. Microbiol.* 1, 395-399

Ringli, C., Baumberger, N., Diet, A., Frey, B., & Keller, B. (2002) ACTIN2 is essential for bulge site selection and tip growth during root hair development of Arabidopsis. Plant Physiol. 129, 1464-1472

Schaefer, D., Zrÿd, J-P., Knight, C.D., & Cove, D.J. (1991) Stable transformation of the moss *Physcomitrella patens*. *Mol. Gen. Genet.* 226, 418-424

Schaefer, D.G. (2002) A new moss genetics: Targeted mutagenesis in *Physcomitrella patens*. *Annu. Rev. Plant Biol.* 53, 477-501

Schaefer, D.G., & Zrÿd, J-P. (1997) Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J.* 11, 1195-1206

Schiestl, R.H., & Petes, T.D. (1991) Integration of DNA fragments by illegitimate recombination in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 88, 7585-7589

Schmidt, W.M., & Mueller, M.W. (1999) CapSelect: a highly sensitive method for 5' CAP-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. *Nuc. Acids Res.* 27, e31

Schumaker, K.S., & Dietrich, M.A. (1998) Hormone-induced signaling during moss development. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49, 501-523

Seifert, H.S., Chen, E.Y., So, M., & Heffron, F. (1986) Shuttle mutagenesis: A method of transposon mutagenesis for *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 83, 735-739

Seifert, H.S., So, M., & Heffron, F. (1986) Shuttle mutagenesis: A method of introducing transposons into transformable organisms. In: Setlow JK, Hollaender A (eds) *Genetic Engineering: Principles and methods,* (pp. 123-133) Plenum Press, New York

Seki, M., Carninci, P., Nishiyama, Y., Hayashizaki, Y., & Shinozaki, K. (1998) High-efficiency cloning of *Arabidopsis* full-length cDNA by biotinylated CAP trapper. *Plant J.* 15, 707-720

Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M., Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., & Shinozaki, K. (2002) Functional annotation of a full-length *Arabidopsis* cDNA collection. *Science* 296, 141-145

Springer, P.S. (2002) Gene traps: Tools for plant development and genomics. *Plant Cell* 12, 1007-1020

Stapleton, M., Liao, G., Brokstein, P., Hong, L., Carninci, P., Shiraki, T., Hayashizaki, Y., Champe, M., Pacleb, J., Wan, K., Yu, C., Carlson, J., George, R., Celniker, S., & Rubin, G.M. (2002) The *Drosophila* gene collection: identification of putative full-length cDNAs for 70% of *D. melanogaster* genes. *Genome Res.* 12, 1294-1300

Sundaresan, V,. Springer, P., Volpe, T., Haward, S., Jones, J.D.G., Dean, C., Ma, H., & Martienssen, R. (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes Dev.* 9, 1797-1810

Suzuki, H., Fukunishi, Y., Kagawa, I., Saito, R., Oda, H., Endo, T., Kondo, S., Bono, H., Okazaki, Y., & Hayashizaki, Y. (2001) Protein-protein interaction panel using mouse full-length cDNAs. *Genome Res.* 11, 1758-1765

Suzuki, Y., Yamashita, R., Nakai, K., & Sugano, S. (2002) DBTSS: DataBase of human Transcriptional Start Sites and full-length cDNAs. *Nuc. Acids Res.* 30, 328-331

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796-815

Walden, R. (2002) T-DNA tagging in a genomics era. *Crit. Rev. Plant Sci.* 21, 143-165

Weigel, D., Ahn, J., Blazquez, M., Borevitz, J., Christensen, S., Fankhauser, C., Ferrandiz, C., Kardailsky, I., Malancharuvil, E., Neff, M., Nguyen, J., Sato, S., Wang, Z., Xia, Y., Dixon, R., Harrison, M., Lamb, C., Yanofsky, M., & Chory J. (2000) Activation tagging in Arabidopsis. *Plant Physiol.* 122, 1003-1013

LIPID METABOLISM IN MOSSES

KOJI MIKAMI AND ELMAR HARTMANN

Department of Regulation Biology, National Institute for Basic Biology, Okazaki, Japan (KM); Institut für Pflanzenphysiologie und Mireobiologie, Freie Universität Berlin, Germany (EH)

Abstract. Lipid composition in mosses is characterized by the presence of a large amount of neutral lipids and polyunsaturated very-long-chain fatty acids, sucu as arachidonic acid and eicosapentaenoic acid. Such characteristics in lipid composition is thought to be important for survive of mosses under various kinds of environmental stresses. A recent establishment of an Expression Sequence Tag (EST) database provides a novel strategy to analyze the physiological significance of unique lipid composition and mechanisms of the lipid metabolism in mosses, such as *Physcomitrella patens* and *Ceratodon purpureus*. By comparison of the *Physcomitrell*a EST database with the whole genome sequence of *Arabidopsis thaliana*, we found that *Physcomitrell*genes that are involved in the regulation of the lipid metabolism were mostly similar to those in *A. thaliana*, except for genes for the production of storage lipids and the oxylipin metabolism. It remains to elucidate why the lipid compositions of these two species are different, despite the high similarity of genes for the lipid metabolism. Analysis of phototropism in *C. purpureus* has demonstrated the importance of light-dependent activation of phosphatidylinositol-specific phospholipase C (PI-PLC), however almost is unclear to understand how PI-PLC regulates the phototropism. Studies on the molecular bases of above characteristics are necessary to understand the physiological importance of the lipid composition and lipid metabolism in mosses.

1. INTRODUCTION

Bryophytes are phylogenetically placed between algae and pteridophytes. They are divided into three classes (1) liverworts (Hepaticophytina); (2) mosses (Bryophytina); (3) hornworts (Anthocerophytina). The bryophytes in spite of having a great deal in common with the green algae progressively developed the vegetative and generative markings. This is understood as an enhanced development depending on the adaptation to the life on land. The mosses conquered the dry land and settled it with the dominating amount of species. The adaptation to this biosphere is very broad and diverse, especially to extreme conditions like aridity and cold regions. This is one of the reasons that the Bryophytina dominate the primary production by plants in regions of high polar latitude and alpine mountain habitats, where the productivity of vascular plants is much reduced. In the Russian tundra about 50% of the biomass is represented by mosses (Aro and Karunen 1979), but also in the arctic coniferous forests of Finland the biomass production of mosses

A.J. Wood et al. (ed.),
New Frontiers in Bryology: Physiology, Molecular Biology and Functional Genomics, 133–155.

correspond to 10-20% of the growth rate of the trees (Malkönen 1974). In the deciduous forests of lower latitude the amount of mosses on the biomass is only 1% (Forman 1969). The understanding of the biochemistry and the physiology of mosses is generally much less developed and there is only fragmentary information about the nature of metabolic regulations and controls. Many reason for these stringent adaptation profiles of mosses were discussed, but no convincing concept that could take into account of all phenomena is so far known.

There is a possibility that the modulation of lipids may have an important regulatory function on the adaptation capacities of mosses, because lipids are very important and effective storage compounds and the main constitutional compounds of cell membranes as glyco- and phospholipids. Bryophytes, for example, contain higher amounts of neutral lipids in green tissue than most other plants (Dembitsky 1993, Jamieson and Reid 1975, Zhukova and Aizdaicher 1995). The most abundant fatty acids of the bryophytes are also common to most other organisms. However, by compared to higher plants, many members of moss, algae and fern families produce a wide variety of polyunsaturated very-long-chain fatty acids (PUVLCFA´s) such as arachidonic acid (AA) and eicosapentaenoic acid (EPA) which are produced only by lower plants (Dembitsky 1993, Jamieson amd Reid 1975, Zhukova and Aizdaicher 1995). We suspect that the unique lipid composition is important for the ability of many bryophytes to survive harsh environmental conditions. The study of lipids in bryophyte is therefore in many respects of particular interest, however the knowledge about lipid metabolism of bryophytes is very rudimentary. The study of the fatty acid (FA) patterns of some types of mutants should help to overcome this ignorance. It was observed, for instance, that considerable changes in many fatty acids occurred in senescent or dark-grown material, including changes in the proportion of C_{20} polyenoic fatty acids. Although such differences in fatty acid patterns have been found between strains, these alone are probably not responsible for their physiological differences (Grimsely et al. 1979).

The life cycle of mosses, *Ceratodon purpureus* and *Physcomitrella patens,* is predominantly haploid (Ashton and Cove 1977), which simplifies the phenotypic analysis after mutation or transformation (Schaefer et al. 1991). Indeed, many mutants showing altered phenotypes have been obtained by using chemical treatment of spores or UV-treatment of protonema (for example Ashton and Cove 1977, Lamparter et al 1996). In addition, protoplasts may be isolated from protonemal tip cells, fused with other cells, and regenerated to give somatic hybrids. These hybrids can be used for biomass production in bioreactors to study the biochemistry of wild types compared with mutants or hybrids. Moreover, genes of *P. patens* can be specifically inactivated by gene-targeting (Schaefer 2002), as shown by Schaefer and Zryd (1997) who demonstrated that integration of homologous DNA into the genome of *P. patens* takes place by homologous recombination with a relative efficiency of more than 90% among transgenic plants. Therefore, *P. patens* and *C. purpureus* are excellent tools for the study of the influence of various experimental conditions on development and metabolism (Ashton and Cove 1977).

In this review, we summarize the knowledge of physiological importance and characteristics of the lipid metabolism in mosses, especially *C. purpureus* and *P.*

patens. Moreover, to understand molecular system of the lipid metabolism, we surveyed genes involved in the lipid metabolism based on the EST database of *P. patens*, which allow us to find unique and common characteristics of the lipid metabolism in mosses in comparison with those in higher plants.

2. UNIQUE CHARACTERISTICS OF LIPID METABOLISM IN MOSSES

2.1. Production of Unique Lipids

The lipid composition of bryophytes is quite unique and is not uniform at all. Many characteristic variations exist in the patterns of esterified fatty acids in different bryophyte species and there is evidence that the fatty acid composition is of taxonomic significance (Kohn et al. 1987a,b,c). In higher plants, acetylenic compounds are often characteristic for certain taxa and therefore their presence or absence can be of chemotaxonomic significance (Pohl and Wagner 1972a,b). Although numerous compounds with acetylenic bonds have been isolated from plants and microorganisms (Bychek 1994), acetylenic fatty acids in bryophytes were mainly identified in mosses (Anderson et al. 1974, 1975). Indeed, triglyceride fraction of a few mosses contains acetylenic fatty acids such as 9,12,15-octadecatrien-6-ynoic acid (Anderson *et al.* 1974), 9,12-octadecadien-6-ynoic acid (Anderson et al. 1975) and 11,14-eicosadien-8-ynoic acid (Jamieson and Reid 1976). The reports of several authors (Swanson et al. 1976, Karunen 1981, Ichikawa et al. 1983) and us indicate that acetylenic fatty acids are in fact abundant in the *Dicranaceae* and related families (Kohn et al. 1987b). There is not much information about acetylenic fatty acids from liverworts, however a new acetylenic fatty acid, 9-octadecen-6-ynoic acid, has been identified in the thalli of the liverwort *Riccia fluitans* L. (Kohn et al. 1987a). Based on these findings, we conclude that the occurrence of acetylenic fatty acids is restricted to only a few strains in bryophytes.

The betaine ether-linked glycerolipids belong to another naturally occurring new group of lipids, which were identified in protozoa, bryophytes, higher fungi and higher plants (Dembitsky 1996). Betaine-lipids such as diacylglyceryltrimethylhomoserine (DGTS) are components of biological membranes like other polar lipid classes. The levels of DGTS in moss species were varied from 7 to 14%, which is lower that those (from 15 to 36%) in liverworts (Dembitsky 1996).

2.2. Production of Polyunsaturated Very-Long-Chain Fatty Acids

It is well known that *C. purpureus* and *P. patens* contain considerable amounts of PUVLCFAs, such as AA (20:4 omega 6) and EPA (20:5 omega 3) (Gellerman et al. 1975, Grimsley et al. 1981, Vandekerkhove et al. 1984, Girke et al. 1998, Zank et al. 2002) that play a key role in eicosanoid metabolism in human (Samuelsson 1983). These are found especially in phosphatidylethanolamine (PE) (Hartmann et al. 1986, Euler 1987), and these metabolites seem to be involved in stress regulation (Vanderkerkhove et al. 1984), representing yet another potential source for second messengers (Anderson 1989). In higher plants, polyunsaturation of VLCFAs resulted in perturbation of membrane structure in chloroplasts and alteration in morphology (Millar et al. 1998), thus VLCFAs are not polyunsaturated generally in higher plants. Therefore, physiological roles of polyunsaturated

PUVLCFAs are important questions to understand physiological significance of the lipid metabolism in moss plants.

The biosynthesis of AA and EPA generally starts with linoleic acid (18:2), which is synthesized in a widely branched network of desaturation and elongation steps (Shiran et al. 1996). The Δ5- and Δ6-desaturases are the key enzymes in this network, which introduce the new double bond between the first double bond and the carboxyl terminus of the fatty acid, known as carboxyl-directed desaturation. This mode differs from the methyl-directed desaturation, which works towards the methyl end of the unsaturated fatty acid. Desaturases of both types belong to the membrane-bound desaturases, which operate in microsomes or in plastids (Heinz 1993). All desaturases, including acyl-ACP, (Ohlrogge and Jawoeski 1993), acyl-CoA (Enoch et al. 1976) and acyl-lipid desaturases, are believed to catalyse an O_2-dependent reaction, in which either cytochrome b_5 serves as electron donor for the microsomal or ferredoxin for the plastidial desaturases (Kearns et al. 1991, Schmidt and Heinz 1990, Smith et al. 1990).

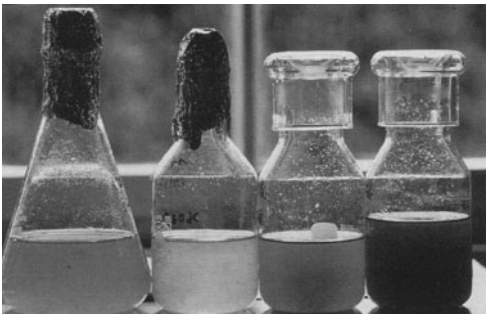

Figure 1. Batch culture of Ceratodon purpureus *in white light (70 W/m^{-2}; 16h light/8 h dark; 20°C). The colour changes form deep green to brownish during the culture of 21 days. Cells of this type are full of oil drops like in Figure 2.* *(See colour section page xi)*

In the last few years, extensive sequence information from various desaturases in the methyl-directed group has been accumulated, but only a few from the carboxyl-directed group have been cloned so far (Reddy et al. 1993, Sayanova et al. 1997). Lipids of *P. patens* contain high proportions of AA (up to 30% of total fatty acids) indicating strong expression of Δ5- and Δ6-desaturases (Grimsley et al. 1981). *P. patens* synthesizes PUVLCFAs by sequential modifications of linoleic acid (18:2) and α-linolenic acid (18:3) by Δ3-desaturase, Δ6-desaturase, Δ6-elongase, and Δ5-desaturase. Genes encoding Δ6-desaturase and Δ6-elongase were recently cloned and analyzed by targeted gene disruption by homologous recombination (Girke et al. 1998, Zank et al. 2002). By disruption of the gene for Δ6-desaturase, arachidonic acid (20:4) was completely loss in the mutant cells, however there was no obvious phenotype (Girke et al. 1998). Moreover, disruption of the gene for Δ6-elongase resulted in a complete loss of all C20-polyunsaturated PUVLCFAs including AA (20:4) and EPA (20:5). Again, no visible alteration in phenotype was observed under normal conditions and stress conditions such as low-temperature, high salinity and drought (Zank et al. 2002). These experimental data clearly indicates that C20-PUVLCFAs are not essential for viability of *P. patens.*

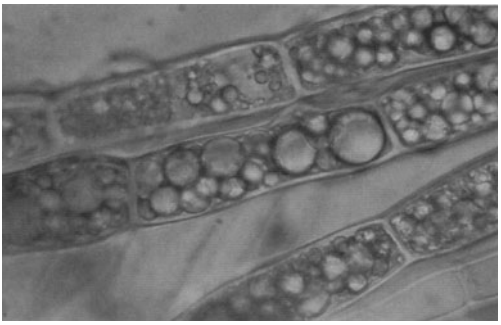

Figure 2. Senescent protonema filaments of Ceratodon purpureus *from 21 day-old batch culture. The cells are full of oil drops. (See colour section page xi)*

Table 1. Modulation of pigment content and the amounts of acylglyerolipids during batch culture of Ceratodon purpureus *in white light (70 W/m^{-2}; 16 h light/8 h dark day; 20°C).*

Culture days	0	3	6	8	10	12	21	84
Growth rate mg dry weight/litre	70.0	115.3	254.0	454.0	570.0	667.9	1012.0	1704.8
Chlorophyll a/b mg/g dry weight	16.8	17.9	15.6	12.6	9.4	7.9	4.6	1.8
Carotenoids mg/g dry weight	3.6	4.7	3.4	3.1	2.7	2.6	1.8	1.3
Phospholipids nmolFA/mg dry weight	36.2	32.6	27.9	24.1	21.8	13.5	9.2	3.8
Glycolipids nmolFA/mg dry weight	46.1	78.3	82.6	47.2	34.5	21.5	15.0	8.0
Total fatty acid content nmol/mg dry weight	101.8	133.5	147.8	117.5	125.2	188.6	247.6	496.5

P. patens has multiple genes for Δ6-desaturase, while Δ6-elongase is encoded in a single copy gene (see Table 2). In fact, *Arabidopsis thaliana* also contains multi-copy of genes for Δ6-desaturase and a single copy of a gene for Δ6-elongase. Because PUVLCFAs were not observed in microsomes in higher plants, it is still question whether Δ6-desaturase and Δ6-elongase are active in *A. thaliana* and, if so, why *A. thaliana* needs polyunsaturated VLCFAs.

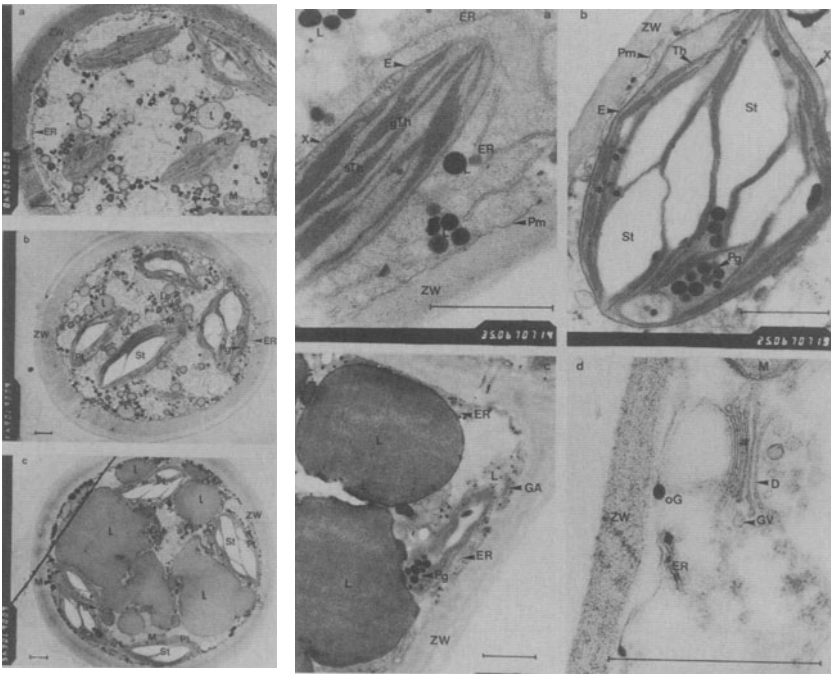

Figure 3 (left). Sequence of transversal section from protonema cells of batch cultured Ceratodon purpureus *shown in Figure 1. 3a, 300-400 mg dry weight/litre medium (bar: 8000x); 3b, 600-800 mg dry weight/litre medium (bar: 6000x); 3c, >100 mg dry weight/litre medium (bar 6000x). Abbreviation: ER, Endoplasmic reticulum,; L, lipid drops; M, mitochondria; Pg, plastoglobuli; PL, chloroplast; St, starch; ZW, cell wall.*

Figure 4 (right). Magnification of the cells showed in Figure 3. 4a, Chloroplast of a young cell showed in 3a (bar 35000x); 4b Chloroplast of an older cell showed in 3b (bar 25000x); 4c, Lipid drop from an old cell showed in 3c (bar 15000x); 4d, Golgi apparatus from an old cell showed in 3c (bar 60000x). Abbreviation: E, chloroplast envelop; ER, Endoplasmic reticulum; D, dictyosome; GA, Golgi apparatus; GV, Golgi vesicle; Th, thylakoids; gTh, grana-thylakoids; sTH, stroma thylakoids; St, starch; oG, osmophilic granula; Pm, plasmamembrane; Pg, plastoglobuli; X, invagination of inner chloroplast membrane; ZW, cell wall.

2.3. Changes in Lipid Composition During Growth

Many bryophytes can easily be kept in axenic culture (Figure 1). The changes in the lipid- and fatty acid pattern of *C. purpureus* during a batch culture reflect the modulation in the lipid metabolism (Figure 1 and Figure 2). In the young well-growing protonemal culture, the biosyntheses of structural lipids is preferred, whereas in the later stage of culture the fatty acid synthesis is predominant (Table 1). This stage of growth is marked by the highest amounts of structural lipids in the plastids (MGDG, DGDG, SQDG and PG) and these lipids were degraded during continuing batch growth (Table 1). With increasing senescence, the triglycerides accumulate and represents at the end of the culturing period up to 15% of dry matter (Figure 2 and Figure 3). The protonema are full alive, the fatty acids are degraded and converted to carbohydrates by anabolic metabolism after inoculation to fresh medium.

Besides triglycerides, structural lipids of none plastidic source, such as phosphatidylserine (PC) and PE, are synthesized. Senescence of mosses is always accompanied by increased in storage lipids. Old protonemas also increase the amount of tmema (dead separating cells), which make the vegetative propagation of protonema pieces easier. Lipids are also the main storage compounds in spores (Huneck 1983, Karunen 1972, Karunen and Liljenberg 1978).

The changes in lipid metabolism can be shown by electronmicroscopic studies of the senescing protonema culture (Figure 3 and Figure 4). The physiological alteration in the lipid metabolism modifies the structure of the plastids and increases the amount and the size of the oil droplets in the cytosol. The thylakoids are much reduced between the enlarged starch grains. Osmiphilic lipid drops develop in great numbers. These metabolic reactions are further accelerated by the senescing process of old cultures where the oil drops fuse to huge drops (Figure 2 and Figure 4).

3. GENES INVOLVED IN THE LIPID METABOLISM IN MOSSES

3.1. Overview of the Lipid Metabolism in Plants

Glycerolipids in plant membranes are primarily composed of 16-carbone and 18-carbon fatty acids containing up to three double bonds. Production of glycerolipids consists of *de novo* fatty acid synthesis and incorporation of fatty acids into glycerol-3-phosphate (G3P) (Ohlrogge and Jaworski 1997, Wallis and Browse 2002), which are represented in Figure 5. Biosynthesis of fatty acids occurs in the plastid where acetate (C2) is elongated by the sequential addition of further C2 units while attached to a soluble acyl-carrier protein (ACP). The glowing acyl chain is terminated when it is 16 or 18 carbon long, which resulted in making 16:0-ACP and 18:0-ACP. Synthesis of common monounsaturated fatty acids is catalyzed by a soluble plastidial desaturase (FAB2 in *Arabidopsis thaliana*), which introduces a double bond between carbons 9 and 10 of a C18 acyl-ACP (□9 position, counting from the carboxyl end) to make 18:1-ACP (oleoyl-ACP).

Incorporation of fatty acids into G3P is catalyzed in both plastids and Endplastic reticulam (ER) (Figure 6). To enter plastidial pathway for synthesis of glycerolipids (so called prokaryotic pathway), fatty acids are directly transferred from ACP to G3P by glycerol-3-phosphate acyltransferase (GPAT), which results in

the production of phosphatidic acid (PA). It is well known that the substrate specificity of plastidial GPAT is different between chilling-tolerant and chilling-sensitive plants (Murata and Tasaka 1997). GPAT in chilling-tolerant plants makes PAs containing 16:0 at the *sn-2* position and 18:1 at the *sn-1* position, while PAs that is made by GPAT in chilling-sensitive plants has 16:0 at both *sn-1* and *sn-2* positions. Because fatty acid profiles of moss plants indicated the existence of polyunsaturated 16- and 18-carbon fatty acids (Vandekerkhove et al 1984, Zank et al. 2002), mosses are thought to have the chilling-tolerant type of GPAT.

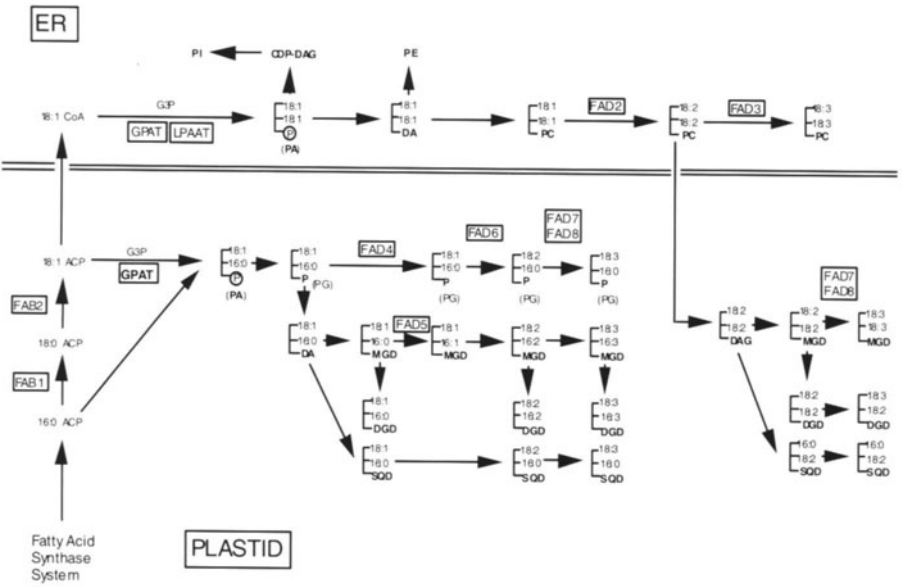

Figure 5. The pathway of glycerolipid synthesis in plants. This figure is based on the pathway in Arabidopsis *leaves (Wallis & Browse, 2002). Abbreviations: G3P, glycerol-3-phosphate; PA, phosphatidic acid; DA, diacyl; CDP-DAG, cytidine-5-diphosphate-diacylglycerol; PG, phosphatidylglycerol;; MGD, monogalactosyldiacyl; DGD, digalactosyldiacyl; SQD, sulfoguinovosyldiacyl; PC, phosphatidylcholine; PI, phosphatidylinositol; PE, phosphatidylethanolamine; GPAT, glycerol-3-phosphate acyltransferase; LPAAT, lysophosphatidic acid acyltransferase; ER, Enoplastic reticulum.*

Once PA is made, it is used for synthesis of phosphatidylglycerol (PG) that is converted to diacylglycerol (DAG) by a PA-phosphatase (PAP). DAG is further converted to other major plastid membrane lipids, monogalactosyl diacylglycerol (MGDG), digalactosyl diacylglycerol (DGDG), and sulfoquinovosyl diacylglycerol (SQDG). Genetic and molecular biological approaches were succeeded to identify desaturases for these glycerolipids, termed FAD for fatty acid desaturation (Figure 5) (Wallis and Browse 2002). FAD proteins are members of a family of membrane-bound acyl-lipid desaturases. Among them, FAD4 and FAD5, which are Δ4 and Δ7 desaturases, have substrate specificity to 16:0 at the *sn-2* position of PG and MGDG, respectively. In contrast, FAD6 for Δ12 desaturase and FAD7 and FADS both for

ω3 desaturase (ω3 position, counting from the N-terminal end) can introduce double bond at the *sn-1* position of PG, MGDG, DGDG and SQDG.

To enter the pathway in ER (so called eukaryotic pathway), in most cases, an acyl-ACP thioesterase cleaves 18:1-ACP into the CP and a free fatty acid. Free fatty acids are exported to the cytoplasm and esterified to CoA by Acyl-CoA synthetase, and then 18:l-CoA moves to microsomes to be used for making glycerolipids. PA is synthesized by two sequential esterification of G3P with 18:l-CoA at the *sn-1* position by GPAT to produce lysophosphatidic acid (LPA) and then at the *sn-2* position of LPA by lysophosphatidic acid acyltransferase (LPAAT) in the eukaryotic pathway. Subsequently, PAP yields both DAG, which is converted to phosphstidylcholine (PC) by choline phosphotransferase (CPT), and CDP-DAG that is a substrate of phosphatidylinositol synthase (PIS) for producing phosphatidylinositol (PI). Fatty acids that are esterfied to PC at the *sn-1* and *sn-2* positions are then unsaturated with microsomal membrane-bound acyl-lipid desaturases, such as FAD2 and FAD3 for Δ12 and ω3 desaturases (Figure 5).

3.2. Similarity of the Overall Strategy in the Lipid metabolism between **Physcomitrella patens** *and* **Arabidopsis thaliana**

Rapid increase of information of ESTs of *P. patens* enables us to compare the genes for lipid metabolism in *P. patens* with those of the genomes of *A. thaliana* and rice. We compared EST clones encoding enzymes involved in the lipid metabolism in *P. patens* with those from genome information of *A. thaliana.* Survey of EST clones of *P. patens* was essentially performed by BLAST search using sequences of *Arabidopsis* genes listed in the Lipid Gene Catalogue (http://www.plantbiology.msu.edu/gene_survey/front_page.htm) against the database of *Physcomitrella* EST program (PEP; http://www.moss.leeds.ac.uk/). Searching of some of genes in *P.a patens* was performed with genes from plants other than *A. thaliana* or animals.

As shown in Table 2, *Physcomitrella* homologues of genes whose products have been annotated as enzymes involved in lipid metabolism in *A. thaliana* were found for most of all genes listed in the Lipid Gene Catalogue for *Arabidopsis*. Therefore, outline of the lipid metabolism in *P. patens* is very similar to that in *A. thaliana*. One example is a synthetic pathway for PE. We have already cloned the *Physcomitrella* cDNA encoding a phosphatidylserine decarboxylase (PSD) (von Orlow, 2000), which catalyses the synthesys of PE from PS (Voelker 1997, see Figure 6), and found a highly conserved structure among PSDs from various species from bacteria to animals and higher plants (Figure 6). Since Table 2 indicates the existence of PS synthase in *P. patens*, it seems that the pathway to produce PE from serine is conserved between *P. patens* and higher plants.

3.3. Differences in Lipid Metabolisms between **Physcomitrella patens** *and* **Arabidopsis thaliana**

Despite similarity in the overall strategy of the lipid metabolism, significant differences were observed for genes involved in synthesis of storage lipids and lipid signalling between *Physcomitrella* and *Arabidopsis* plants (Table 2).

This probably reveals the differences in the details the lipid metabolism. We here mention two examples for such differences between these two plants.

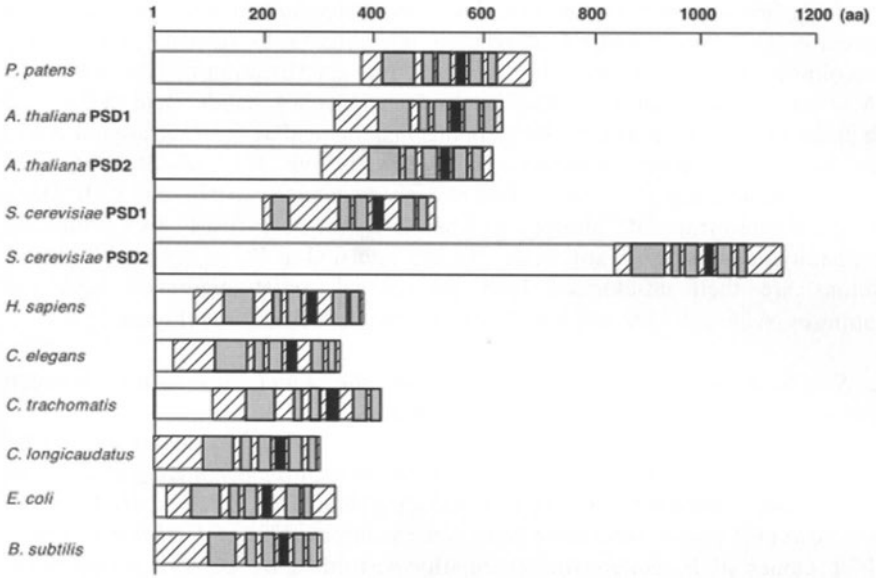

Figure 6. Comparison of phosphatidylserine decarboyxlase (PSD) from different organism. The PSD is first synthesized as a precursor which contains α and β subunits, then matured by endoproteolytic cleavage. The region containing two subunits corresponds to the C-terminal 300 amino acids of all PSDs (slashed areas). The regions which represents 60 to 100 % identity among PSDs are indicated by gray bars. The region containing LGST- or GGST-motif which is predicted to be the cleavage site for α and β subunits (Voelker, 1997) are indicated by black bars. GenBank accession nos.: A. thaliana *PSD1 and PSD2, CAB39662 and BAA97369, respectively;* S. cerevisiae *PSD1 and PSD2, AAA34918 and AAA69819, respectively*; H. sapiens, *AAH09315*; C. elegans, *NP_498601*; C. trachomatis, *AAC68294;* C. longicaudatus, *CAA56630;* E. coli, *AAC77120;* B. subtilis, *BAA07226.*

3.3.1. Production of Storage Lipids

The most obvious difference was found for EST clones encoding oil-body oleosin, caleosin and acyl-CoA:diacylglycerol acyltransferase (Table 2) that are involved in synthesis of storage lipids. In general, plants accumulate triacylglycerol (TAG) as a storage lipid to regulate cellular metabolism and to serve as carbon and energy stores (Voelker and Kinney 2001). TAG is indeed present in most plant organs, including leaves, petals, fruits, anthers, and developing seeds (Hobbs et al. 1999). Synthetic patheway of TAG was well studied in seeds in which TAG is first synthesized in the membranes of the ER and then released into the cytosol to form the oil bodies as lipid storage organelles. Oil bodies have a matrix of TAG that is surrounded by a monolayer of phospholipid associated with oleosins and caleosins, both of which are the abundant structural protein and a newly identified minor

constituent, respectively (Frandsen et al. 2001). Because *P. patens* does not produce seeds, it is reasonable that there are no genes for oleosin and caleosin.

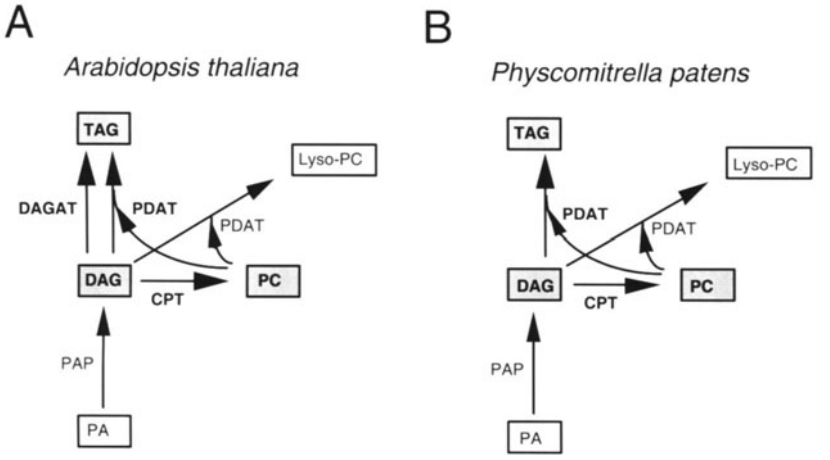

Figure 7. Schematic representation of proposal pathways for triacylglycerol production. Survey of Physcomitrella *ESTs clearly indicated the lack of DAGAT in this organism, thus the pathway of TAG formation seems to be different between* Arabidopsis thaliana *and* Physcomitrella patens. *A,* Arabidopsis thaliana*; B,* Physcomitrella patens. *Abbreviations: TAG, triacylglycerol; DAG, diacylglycerol; PC, phosphatidylcholine; PA, phosphatidic acid; DAGAT, acyl-CoA:diacylglycerol acyltransferase; PDAT, phosphatidylcholine:diacylglycerol acyltransferase. CPT, choline phosphotransferase.*

The absence of acyl-CoA:diacylglycerol acyltransferase (DAGAT) and the presence of phosphatidylcholine:diacylglycerol acyltransferase (PDAT) (Table 2) are important information to understand how *Physcomitrella patens* produces TAG. DAGAT, a final enzyme for TAG synthesis, is capable of transferring an acyl group from acyl-CoA to the *sn-3* chain of DAG to form TAG (Figure 7a). PDAT can transfer the *sn-2* acyl chain from PC, a major component of membranes, to DAG to form TAG and lyso-PC (Dahlqvist et al. 2000). Moreover, choline phosphotransferase (CPT), which esterifies *sn-3* of DAG with choline to produce PC, is necessary for the PDAT pathway (Voelker and Kinney 2001). The absence of DAGAT indicates that *P. patens* produces TAG via only PDAT pathway (Figure 7b), which is supported by the presence of genes for CPT in *P. patens* (Table 2).

3.3.2. Oxylipin Metabolism

Biosynthesis of oxylipins rises from α-linolenic acid (LA) to produce aldehyde, ketol, epoxy, hydroxy, and divinyl-ester derivatives (Figure 7) (Howe and Schilmiller 2002, Feussner and Wasternack 2002). One of the routes for oxylipin-biosynthesis pathways is known as the octadecanoid pathway in which jasmonic acid (JA) and methyl jasmonate (MeJA) are finally produced (Turner et al. 2002). For biosynthesis of oxylipins, LA is first released from membranes by

phospholipase A (PLA) and then added oxygen to the 9 or 13 position of the C18 chain by lipoxyganase (LOX) (Figure 8).

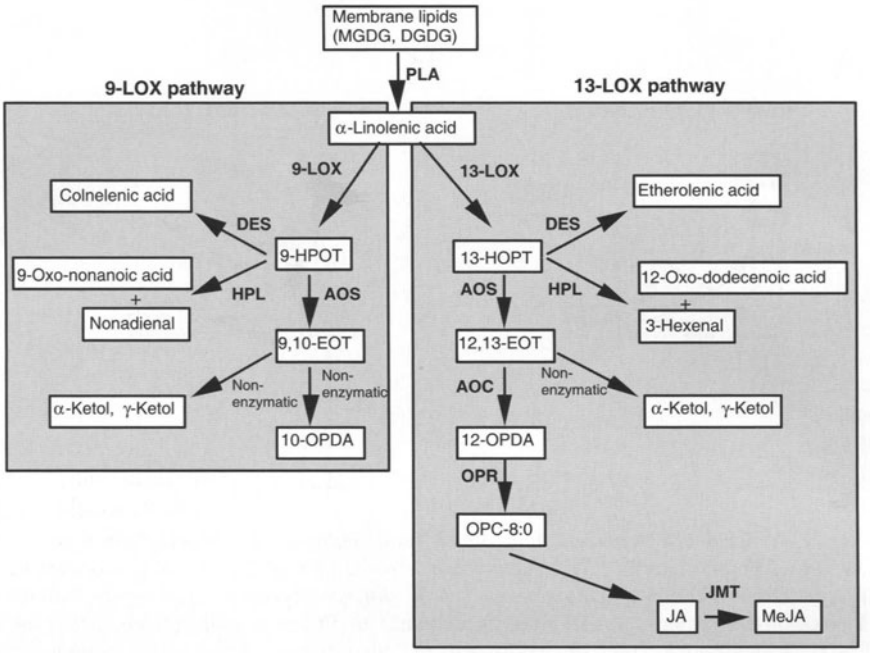

Figure 8. The 9-LOX and 13-LOX pathways in oxylipin biosynthesis. In chloroplasts, α-linolenic acid is released from membranes with PLA and metabolized to produce signalling compounds, such as OPDAand JA, by lipoxygenases in both 9-LOX and 13-LOX psthways. The 9-LOX pathway is completed in chloroplasts. In the 13-LOX psthway, OPDA is released to cytoplasm and converted to OPC-8:0 which is further translocated into microbody to generate JA and MeJA. Abbreviations: LOX, lipoxygenase; DES, divinyl ester synthase; AOS, allen oxide synthase; AOC, allen oxide cyclase; HPL, hydroperoxide lyase;, OPR, 12-oxo-phytodienoic acid reductase; HPOT,, hydroperoxy octadecatrienoic acid; EOT, epoxy-octadecatrienoic acid; OPC-8:0, 3-oxo-2-cyclopentane-1-octaoic acid; JA, jasmonic acid; MeJA, methyl jasmonate; JMT, JA carboxyl methyltransferase.

PLA family consists of two subfamilies, PLA1 and PLA2, which hydrolyse the *sn-1* and *sn-2* acylester bond of phospholipids to release fatty acids from membranes (Wang 2001). Major targets of PLA are monogalactosyl diacylglycerol (MGDG) and digalactosyl diacylglycerol (DGDG) in chloroplasts. It was reported that both PLA1 and PLA2 are involved in oxylipin biosynthesis (Dhondt et al. 2000, Ishiguro et al. 2001). LOXs are also divided into 9-LOX and 13-LOX that generate 9- and 13-hydroperoxy octadecatrienoic acids (9- and 13-HPOT), respectively, to lead the 9-LOX and 13-LOX pathways of oxylipin biosynthesis (Howe and Schilmiller 2002, Feussner and Wasternack 2002, Turner et al. 2002). In 13-LOX pathway, allene oxide synthase (AOS) converts 13-HPOT to 12,13-epoxy-octadecatrienoic acid (12,13-EOT). Then, 12,13-EOT is cyclized by allene oxide cyclase (AOC) to produce 12-oxo-phytodeinoic acid (12-OPDA) that is

further transformed to JA and MeJA *via* some enzymatic steps including OPDA-reductase (OPR) and JA carboxyl methyltransferase (JMT). 13-HPOT is also converted to other oxylipins, such as etherolenic acid by divinyl ester synthase (DES), 12-oxo-dodecenoic acid and 3- hexenal by hydroperoxidelyase (HPL), and ketols by non-enzymatic processes (Figure 8). In 9-LOX pathway, 9-HPOT is converted to 9,10-EOT by AOS, colnelenic acid by DES, 9-oxo-nonanoic acid and nonadienal by HPL, and ketols by non-enzymatic processes. In contrast to 13-LOX pathway, 10- OPDA produced from 9,10-EOT without enzyme is not converted to JA and MeJA (Figure 8). *P. patens* has genes for all enzymes involved in biosynthesis of oxylipins, except for JMT.

We found 10, 15 and 9 genes for HPL, AOS and DES, respectively, in *P. patens* (Table 2), however the whole sequence of the *Arabidopsis* genome revealed the presence of a single gene for HPL and AOS and no gene for DES. In fact, differences in the existence of genes for genes for LOX, AOS, DES and HPL have already compared among *A. thaliana, Lycopersicon esculentum* (tomato), *Nicotiana attenuata*, and *Mediago sativa* (see Howe & Schilmiller, 2002). Of these plants, only tomato is known to have genes for LOX, AOS, DES and HPL (Howe and Schilmiller 2002), thus it is rare case for the existence of all these genes for the oxylipin biosynthesis in a genome as is in *P. patens*. It has been indicated that oxylipin biosynthesis is regulated by wounding, pathogen attack, touch, osmotic shock, drought and UV light, and during development (Howe and Schilmiller 2002). Moreover, 12-OPDA itself can act as a signal molecule that regulates mechanotransduction and wound-inducible genes (Stelmach *et al.* 1998, Stintzi et al. 2001), while 10-OPDA is involved in tuber formation in potato (Kolomiets et al. 2001). These findings raise the possibility that individual molecules of LOX, AOS, DES, OPR and HPL regulate the biosynthesis of oxylipins under various stress conditions, at different stages of development, and in different cellular compartments in *P. patens*. Alternatively, multiplicity of genes for oxylipin-biosynthetic enzymes may reveal that oxylipins are very important components for growth and survive under stress conditions in *P. patens*. Functional analysis of individual genes and oxtlipins will be essential to understand the species-specific mode of the oxylipin metabolism in *P. patens.*

4. PHYSIOLOGICAL ROLES OF LIPIDS IN MOSSES

4.1. Phototropism in Mosses

Adaptation to the changes in environment is a crucial mechanism in the development of sessile plants. Among environmental stresses, light is the most important for plants (Sopory et al. 1999). Photomorphogenetic reactions are induced, for example, by both red light and blue light. Responses to red light result from perception by the photoreversible pigment, phytochrome, which is conserved in various photosynthetic organisms including cyanobactera, mosses and higher plants (Smith 2000). Recently, the blue-light receptor was identified in some plants (Lin 2002), although it has not yet been identified in mosses. Whereas phytochrome is a well-characterized photoreceptor that mediates different morphogenetic responses, there is little information on the signal transducing mechanisms that

couple light perception to induction of internal morphogenetic processes (Sopory et al. 1999).

Moss protonemata have several advantages as a model system to study light-regulated photomorphogenesis (Hughes and Hartmann 1999). For example, the tip cell of moss protonema (Figure 9) is less differentiated and much simpler morphologically than the complex tissues of higher plants that are frequently employed in studies of photomorphogenesis. This allows direct observation of the morphogenetic response in close correlation with the stimulus by different microscopic and histochemical techniques. The influence of light on the rate and the orientation of moss protonemal cells were first demonstrated by Pringsheim and Pringsheim (1935). Briere and coworkers (1979) demonstrated the influence of light and age on the growth of *Ceratodon* protonema. It was subsequently shown that phytochrome is the photoreceptor for positive phototropism in most fern and all moss protonemata (Etzold 1965, Hartmann et al. 1983, Hartmann and Jenkins 1984, Lamparter et al. 1996,1997), which is in contrast to higher plants. The involvement of phytochrome is unequivocally demonstrated by the action spectra for the response and the reversal (Hartmann *et al.*, 1983). Hartmann and Weber (1988) were able to separate temporarily light perception from the photomorphogenetic growth reaction. The stimulus from unilateral light was stored during growth inhibition by monensin and was expressed as directed growth when growth resumed.

In lower plants like algae and mosses, photomorphogenetic processes seem to be preferably operated *via* modulations of the response. The morphogenetic response is reversed after removal of the inducing factor. In contrast, during differentiation, the morphogenetic change remains even after the inducing factor is removed. In higher plants, genes that regulate differentiation appear to dominate photomorphogenesis. Nevertheless, there are still some questions about the general action of phytochrome.

The crucial steps of the signal transduction pathway of the photoresponse are, however, only beginning to be understood in mosses (Hartman 1984). Growth of the filamentous moss protonemata occurs exclusively at the apical region of the tip cell (Figure 9). Phototropic studies with dark-grown cultures of *Ceratodon* show a negative gravitropic-response and a subsequent bending in response to a unilateral red-light irradiation (Hartmann et al. 1983). When the apical region is irradiated, a swelling and bulging is induced at the tip. This reaction is followed by an orientation of the growth centre toward the irradiated side of the cell (Figure 10).

During the process of phototropism, the accumulation of Ca^{2+} and changes in orientation of microfilaments (MFs) at the apical region of tip protonemal cells were observed (Hartmann and Weber 1988, Meske and Hartmann 1995, Meske et al. 1996), suggesting the involvement of phosphatidylinositol-specific phospholipase C (PI-LC) and phosphatidylinositol-4,5-bisphosphate (PIP_2) in an increase in Ca^{2+} levels (see Figure 11) and regulation of orientation of MFs, respectively. Indeed, when protonemal tissue was labelled with ^{32}P-phosphate in darkness and the radioactivity chased after irradiation with different light programs, there was a clear degradation of PIP_2 and a transient increase in the levels of inositol-l,4,5-trisphosphate (IP_3) and inositol-1,4-bisphosphate (IP_2) (in comparison with steady state amount of IPs of 188±43 pmol/g fresh weight (n=5) in protonemal

cells). Moreover, when dark-adapted protonemata were irradiated and the enzyme activity determined after different irradiation times, the activity of the PI-PLC appeared responsive to control by phytochrome (Hartmann and Pfaffmann 1990). Thus, the phototropism of *Ceratodon* protonemal cells is regulated by phytochrome *via* activation of PI-PLC in red light-dependent manner (Hartmann and Pfaffmann 1990). Phototropic response induced by red light was also observed in protonemal cells of *P. patens* (Jenkins & Cove, 1983), suggesting the involvement of phytochrome and PI-PLC in phototropism in this strain. Although *Physcomitrella* contains at least 7 genes for PI-PLCs (Table 2), it is unclear which isoform of PI-PLCs is involved in the light signalling in *P. patens*.

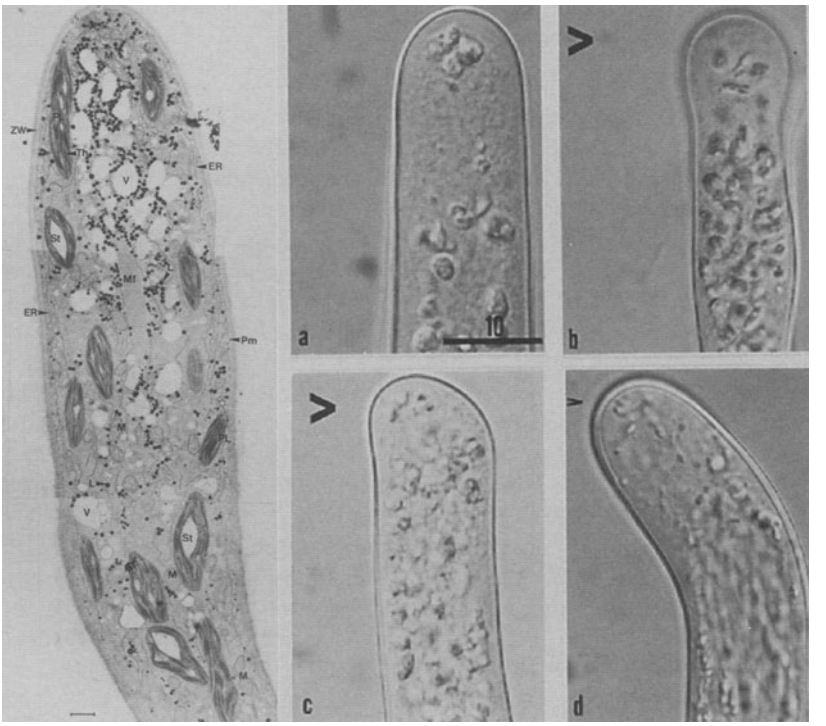

Figure 9 (left). Combined sequence of longitudinal sections of the apical region of the tip cell of the protonema of Ceratodon purpureus *(8 days in white light 16h/d, 20°C; magnification of all sections, bar 8000x). Abbreviations: ER, Endoplasmic reticulum; L, lipid drops; M, motochondria; Mf, microfilaments; PL, chloroplast; Pm, plasmamembrane; St, starch; Th, thylakoids; V, vacuole; ZW, cell wall.*

Figure 10 (right). Response of vertically oriented protonemal tip cell of Ceratodon purpureus *to unilateral red light (660 nm , 6.5 µmol/m^2; bar 10 µm). a, dark control; b, 5 min red light (bulging); c, 20 min red light (displacement of growth point); d, 60 min red light (positive phototropic response).*

Other support for a role of PI turnover in the photomorphogenesis of mosses comes from inhibitor studies. We showed that the monensin-coupled signal-storage in moss protonemata was abolished by applying $5x10^{-3}$ M lithium

(Hartmann and Pfaffmann 1990). The Li^+-effect was expressed best if the lithium was present during and after irradiation. The experiments indicated an influence on the turnover of inositol monophosphate, which accumulated after lithium treatment. This accumulation is caused by an inhibition of the inositol monophosphatase, which interrupts the regeneration of PIP_2. Inhibition of the PI turnover by Li^+ is considered to be the physiological basis for the pharmacological activity of lithium (Huckle and Conn 1987). The inhibition of the PI turnover by lithium was confirmed by Gabler (1988) in beans in connection with the phytochrome-induced hook opening reaction, although no influence of lithium ions on this process was also reported by Morse *et al.* (1987). These above data clearly demonstrates that PI-PLC and breakdown of PI play major roles in light-signal transduction and photomorphogenesis in moss cells.

4.2. Overview of the Phosphatidylinositol Turnover in Plants

Membrane components have a crucial role in translating most hormone- and neurotransmitter-caused signals. Especially, metabolism of phosphoinositides (PI turnover) is now emphasized to play important roles in a variety of cellular processes in both animals and plants (Coté et al. 1996, Munnik et al 1998, Rebecchi and Pentyala 2000, Rhee 2001, Wang 2001). The critical step in the PI turnover is the rapid hydrolysis of PIP_2 by PI-PLC, which results in production of two second messengers, DAG and IP_3 (Gilman 1987, Berridge and Irvine 1984, Munnik et al. 1998, Rebecchi and Pentyala 2000, Rhee 2001, Wang 2001). These second messengers promote the activation of protein kinase C (PKC) and increase in intracellular concentration of Ca^{2+} via activation of IP_3 receptors on Ca^{2+} stores in animals (Streb et al. 1983, Nishizuka 1984, Rebecchi and Pentyala 2000, Rhee 2001). PI turnover is also important in plants (Coté et al. 1996, Munnik et al. 1998, Wang 2001), although PKC and IP_3 receptor have not yet been identified.

PI turnover consists of stepwise phosphorylation of PI with phosphatidylinositol 4-kinase (PI4K) and phosphatidylinositol-4-phosphate 5-kinase (PIP5K) to produce phosphatidylinositol-4-phosphate (PIP) and PIP_2, respectively, and hydrolysis of PIP_2 into DGA and IP_3 by PI-PLC (Munnik et al. 1998, Wang 2001, Müller-Röber and Pical 2002). In plants, since DAG is rapidly converted to phosphatidic acid (PA) by DAG kinase (DGK), PA rather than DAG seems to act as a second messenger in plants (Munnik 2001), which is consistent with the fact that plants lack PKC. Increase in intracellular Ca^{2+} concentration is regulated by not only PLC but also phospholipase D (PLD) that hydrolyzes membrane glycerolipids to synthesize PA and a water-soluble head-group (Wang 2001). It seems that PA produced by PLC and PLD binds to PIP5K and PLC, which accelerates the PI turnover (Figure 11) (Munnik et al. 1998). PI turnover is involved in a variety of physiological processes, especially in responses to light, osmotic stress, and pathogens (Munnik et al. 1998. Wang 2001). Indeed, expression of genes encoding DGK, PLC, PIP5K, and PLD induced by salt, dehydration and ABA treatments (Hirayama et al. 1995, Fan et al. 1997, Katagiri et al. 1996, 2001, Mikami et al. 1998, Müller-Röber and Pical 2002).

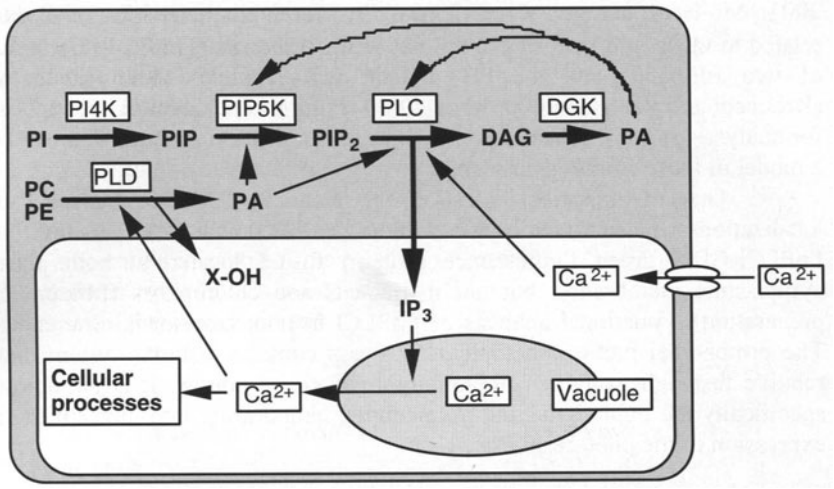

Figure 11. Schematic representation of the relation between PI turnover and generation of second messengers in plants. PLC plays a central role in the production of IP_3 and DAG from PIP_2 which is produced from PI via stepwise phosphorylarion with PI4K and PIP5K. PLD enhances this process by generation of PA which stimulates the activity of PIP5K and PLC. Membranes of vacuole probably contain an IP_3-receptor which regulates the mobilization of intracellular Ca^{2+}. Abbreviations: PC, phosphatidylcholine; PE, phosphatidylethanolamine; X-OH, free head group.

PI is also phosphorylated by phosphatidylinositol 3-kinase (PI3K), resulting in synthesis of phosphatidylinositol 3-phosphate (PI3P) that is further phosphorylated by both PI4K and phosphatidylinositol-3-phosphate 5-kinase (PI3P5K) to produce PI(3,4)P_2 and PI(3,5)P_2, respectively (Vanhaesebroeck et al. 2001). Biological importance of PI(3,5)P_2 in plants was demonstrated by the increase in its concentration under osmostress conditions (Meijer et al. 1999, 2001). Table 2 showed the existence of genes for lipid kinases (PI3K, PI4K, PIP5K, PI3P5K, and DGK) and phospholipases (PLC and PLD) in *P. patens*. Thus, *Physcomitrella patens* appear to be a good experimental system for studies on physiological roles of the PI turnover.

4.3. Phosphatidylinositol Specific-Phospholipase C in P. patens

To elucidate the roles of PI-PLCs in phototropism, we recently isolated a *Physcomitrella* cDNA encoding PI-PLC, designated PpPLC1 (Mikami et al. in preparation). Search of *Physcomitrella* EST bank indicated the multiplicity of PI-PLC (Table 1) and much abundance of cDNA for PpPLC1 in comparison with those for other PI-PLCs. Thus, PpPLC1 is the major isoform among PLCs in *Physcomitrella patens*. Overall structure of PpPLC1 resembles the similarity to those of higher plant PI-PLCs that consist of the N, X, Y, and C2 domains (Figure 12) (Munnik et al. 1989, Wang 2001, Müller-Röber and Pical 2002). The N domain is a variant form of an EF-hand for a Ca^{2+}-binding domain (Otterhag et al. 2001), while the X and Y domains constitute a catalytic domain (Munnik et al. 1989, Wang

2001, Müller-Röber and Pical 2002). Structural characteristics of PpPLC1 are related to those in δ isoforms of animal PLCs, although plant PI-PLCs luck the one of two EF-hands and the PH domain that regulates membrane localization (Rebecchi and Pentyala 2000, Rhee 2001) (Figure 12). Therefore, PpPLCl is useful for analyses of the mechanisms regulating the activation and function of PI-PLC as a model of those from higher plants.

One of important question for plant PI-PLCs is the mechanism of localization at plasma membranes without the PH domain. We already found that PpPLCl-GFP (green fluorescence protein) fusion located at both plasma and cytoplasmic membranes, but not in nucleus and chloroplasts (Mikami et al. in preparation). Functional analysis of PpPLCl in phototropism is now under going. The protonemal moss gametophyte is a less complex cellular system that shows relative fast light triggered photomorphogenetic responses. It is possible to store specifically the light signal and to uncouple, temporarily, light perception from the expression of the photoresponse.

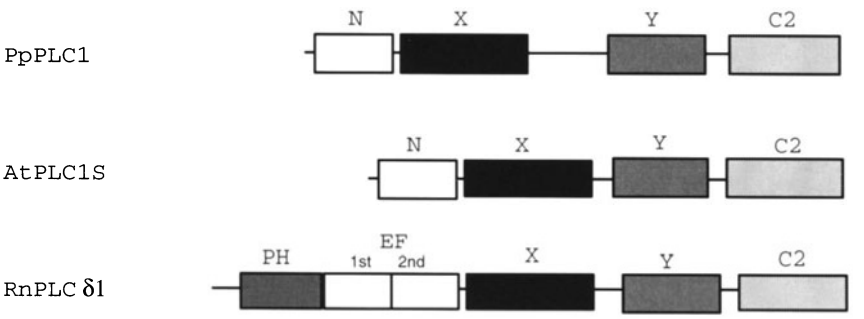

Figure 12. Comparison of the domain organization among PLCs from Physcomitrella patens, Arabidopsis thaliana *and rat. The X, Y and C2 domains are well conserved, whereas the PH domain and one of the EF hands are lacking in plant PLCs. The N domain is a variant of the EF hand that was found in only plant PLCs (Otterhag et al., 2001; Mikami et al., in preparation). GenBank accession nos. of AtPLC1S and RnPLCδ1 are D38544 and NM_017035, respectively.*

The involvement of inositol phosphates *via* a light modulation of PI-PLC in the signal transduction chain was indicated (Hartmann and Pfaffmann 1990). However, it is still too early to be certain that this pathway functions as a specific signaling system for epigenetic light signal transduction. Hopefully, moss protonemata and especially mutant strains are cellular systems that will help us to gain new information about signal amplification and transduction in plants.

5. Outlook

Lipid and fatty acid synthesis is a primary metabolic pathway essential for the function of every plant cell. Its products serve as the central core of membranes in every plant cell, and in specialized cells, fatty acids or fatty acid derivatives act as

signal or hormone molecules, as carbon and energy storage, and as a surface layer protecting the plant from environmental and biological stress. Despite these very diverse functions, all fatty acids in a cell are essentially produced from a single set of enzymes localized mainly in the plastid. Understanding how cells regulate the production of these fatty acids and direct them toward their different functions is thus important for understanding a large range of fundamental questions in plant biology. In addition, much interest has recently developed in genetic engineering of the fatty acid biosynthetic pathway to produce new or improved vegetable oils and industrial chemicals. Therefore, knowledge of how cells control the amount of lipids and fatty acid may be essential for optimal commercial production.

Our understanding of regulation of lipid and fatty acid metabolisms is much less developed than that of carbohydrate or amino acid biosynthetic pathways. What molecules regulate the phospholipids and fatty acid metabolism by feedback or other mechanisms? What metabolic signals or mechanisms control those molecules? How is the global regulation of dozens of genes for lipid synthesis accomplished? We have only fragmentary information about the nature of these controls. Thus, understanding the regulation of the biosynthesis of glycolipids, phospholipids and fatty acids is a relatively unexplored field. The moss systems can play an important role of elucidating those pathways and their physiological background.

6. References

Anderson, M. (1989) Membrane-derived fatty acids as precursors to second messengers. In W.F. Boss and D.J. Morre (Eds), *Second messengers in Plant Growth and development* (pp.181-212). New York: Alan R. Liss, Inc.

Anderson, W.H., Gellerman, J.L. & Schlenk, H. (1975) Acetylenic acids from mosses. Lipids, *10, 501-502*

Anderson, W.H., Hawkins, J.M., Gellerman, J.L., & Schlnk, H. (1974) Fatty acid composition as criterion in taxonomy of mosses. *J. Hattori Bot. Lab.,* 38, 99-103

Aro, E.-M., & Karunen, P. (1979) Effect of changed environmental conditions on glycolipids of the mosses *Pleurozium schreberi* und *Ceratodon purpureus. Physiol. Plant.,* 45, 201-206

Aro, E.-M., Somersalo, S., & Karunen, P. (1987) Membrane lipids in *Ceratodon purpureus* protonemata grown at high and low temperatures. *Physiol. Plant.,* 69, 65-72

Ashton, N.W., & Cove, D.J. (1977) The isolation and preliminary characterisation of auxotrophic and analogue resistant mutants in the moss *Physcomitrella patens. Mol Gen. Genet.,* 154, 87-95

Berridge , M.J., & Irvine, R.F. (1984) Inositol trisphosphate , a novel second messenger in cellular signal transduction. *Nature,* 312, 315-321

Briere, C., Buis, R., & Larpent, J.-P. (1979) Cellular growth and cellular division in relation to age and illumination in the protonema of *Ceratodon purpureus* Brid.. *Z. Pflanzenphys.*, 95, 315-322.

Bychek, I.A. (1994) Pecularities of lipid distribution in Bryophytes: Taxonomic and ecologigal aspects. *Biochemistry (Moscow)* 59, 1227-1238

Coté, G.G., Yueh, Y.G., & Crain, R.C. (1996) Phosphoinositide turnover and its role in plant signal transduction. *Subcell. Biochem.* 26, 317-343

Dembitsky, V.M. (1996) Betaine ether-linked glycerolipids: Chemistry and Biology. *Prog. Lipid Res.,* 35, 1-51

Dembitsky, V.M. (1993) Lipids of bryophytes. *Prog. Lipid Res.,* 32, 281-356

Dhondt, S., Geoffroy, P., Stelmach, B.A., Legrand, M., & Heitz, T. (2000) Soluble phospholipase A2 activity is induced before oxylipin accumulation in tobacco mosaic virus-infected tobacco leaves and is contributed by patatin-like enzymes. *Plant J.,* 23, 431-440

Enoch, H.G., Catalá, A., & Strittmater, P. (1976) Mechanism of rat liver microsomal stearoyl-CoA desaturase. *J. Biol. Chem.,* 251, 5095-5103

Euler, R. (1987) Der Lipid- und Fettsäurestoffwechsel von Protonemazellkulturen des Laubmooses *Leptobryum pyriforme* (Hedw.) Wils. Unter besonderer Berücksichtigung der hochungesättigten C2O-Fettsäuren. PhD Thesis, Johannes Gutenberg Universität Mainz.

Etzold, H. (1965) Der Polrotoropismus und phototropismus der chloronemen von *Dryopterix filix mas* (L) Schott. *Planta,* 64, 254-280

Fan, L., Zang, S., & Wang, X. (1997) Antisense suppression of phospholipase D alpha retards abscisic acid- and ethylene-promoted senescence of postharvested *Arabidopsis* leaves. *Plant Cell,* 9, 2183-2196

Feussner, I., & Wasternack, C. (2002) The lipoxygenase pathway. *Annu. Rev. Plant Biol.* 53, 275-297

Forman, R.T.T. (1969) Comparison of coverage, biomass, and energy as measure of standing crop of bryophytes in various ecosystems. *Bull. Torrey Bot. Club.,* 96, 582-591

Frandsen, G.I., Mundy, J., & Tzen, J.T. (2001) Oil bodies and their associated proteins, oleosin and caleosin. *Physiol. Plant.,* 112, 301-307

Gabler, J. (1988) Lichtinduzierte Signalverarbeitung beim Hypokotylhaken von *Phaseolus vulgaris.* PhD Thesis, Johannes Gutenberg Universität Mainz.

Gellerman, J.L., Anderson, W.H., Richardson, D.G., & Schlenk, H. (1975) Distribution of arachidonic and eicosapentaenoic acid in the lipids of mosses. *Biochim. Biophys, Acta ,* 388, 277-290

Gilman, A.G. (1987) G proteins: Transducers of receptor generated signals. *Annu. Rev. Biochem.,* 56, 615-649

Girke, T., Schmidt, H., Zähringer, U., Reski, R., & Heinz, E. 1998) Identification of a novel D6-acyl-group desaturase by targeted gene disruption in *Physcomitrella patens. Plant J.,* 15, 39-48

Grimsley, N.H., Grimsley, J.M., & Hartmann, E. (1981) Fatty acid composition of mutants of the moss *Physcomitrella patens. Phytochemistry* 20, 1519-1524

Hartmann, E., Klingenberg, B., & Bauer, L. (1983) Phytochrome-mediated phototropism in protonemata of the moss *Ceratodon purpureus* BRID. *Photochem. Photobiol.,* 38, 599-603

Hartmann, E. (1984) Influence of light on phototropic bending of moss protonemata of *Ceratodon purpureus* (Hedw.) Brid. *J. Hattori Bot. Lab.*, 55, 87-98

Harmann, E., & Jenkins, G. (1984) Photomorphogenesis of mosses and liverworts. In A.F. Dyer (Ed.), *The Experimental Biology of Bryophytes* (pp.203-228), London : Academic Press.

Hartmann, E., Beutelmann, P., Vandekerkhove, O., Euler, R., & Kohn, G. (1986) Moss cell cultures as sources of arachidonic and eicosapentaenoic acids. *FEBS Lett.,* 197, 51-55

Hartmann, E., & Weber, M. (1988) Storage of the phytochrome-mediated phototropic stimulus of moss protonemal tip cells. *Planta* 175, 39-49

Hartmann, E. and Pfaffmann, H. (1990) Phosphatidylinositol and phytochrome-mediated phototropism of moss protonemal tip cells. In D.J. Morré, W.F. Boss, & F.A. Loewus (Eds.). *Inositol metabolism in Plants* (pp.259-275). New York: Wiley-Liss, Inc.

Heinz E. (1933) Biosynthesis of polyunsaturated fatty acids. In T.S. Moore, Jr. (Ed.), *Lipid Metabolism in Plants* (pp33-89). Boca Raton, Fl: CRC Press.

Hirayama, T., Ohto, C., Mizoguchi, T., & Shinozaki, K. (1995) A gene encoding a phosphatidylinositol-specific phospholipase C is induced by dehydration and salt stress in *Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA* 92, 3903-3907

Hobbs, D.H., Chaofu, L., & Hills, M.J. (1999) Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. *FEBS Lett.* 452, 145-149

Howe, G.A., & Schilmiller, A. (2002) Oxylipin metabolism in response to stress. *Curr. Opin. Plant Biol.* 5, 230-236

Huckle, W.R., & Conn, P.M. (1987) Use of lithium in measurements of stimulated pituitary inositol phospholipids turnover. *Methods Enzymol.* 141, 149-155

Hughes, J., & Hartmann, E. (1999) Photomorphogenesis in lower plants. In G.S. Singal, G. Renger, S.K. Sopory, K.-D. Irrgang, & Govindjee (Eds.), *Concepts in Photobiology:Photosynthesis and Photomorphogenesis* (pp.834–867). New Delhi, India: Narosa Publishing House.

Huneck, S. (1983) Chemistry and biochemistry of bryophytes. In R.M. Schuster (Ed.), *New Manual of Bryology* (Vol. 1, pp.3-117). Nichinan, Miyazaki, Japan: The Hattori Botanical Laboratory.

Ichikawa, G., Namikawa, M., Yamada, K., Sakai, K., & Kondo, K. (1983) Novel cyclopentenonyl fatty acids from mosses, Dicranum scoparium *and* Dicranum japonicum. Tetrahedon Lett., *24, 3337-3340*

Ishigro, S., Kawai-Oda, A., Ueda, J., Nishida, I., & Okada, K. (2001) The defective in anther dehiscence 1 gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis. Plant Cell,* 13, 2191-2209

Jamieson, G.R., & Reid, E.H. (1975) The fatty acid composition of fern lipids. *Phytochemistry,* 14, 2229-2232

Jamieson, G.R., & Reid, E.H. (1976) Lipids of *Fontinalis antipyretica. Phytochemistry,* 15, 1731-1731

Jenkins, G.I., & Cove, D.J. (1983) Phototropism and polarotropism of primary chloronemata of the moss *Physcomitrella patens. Planta,* 158, 357-364

Karunen, P. (1972) Studies on moss spores. I. The triglycerides of *Polytrichum commune* spores and their mobilization and degradation in relation to the germination phases. *Annals. Univ. Turku Ser., A* 51, 1-70

Karunen, P., & Liljenberg, C. (1978) Content and fatty acid composition of steryl and wax esters in germinating spores of *Polytrichum commune. Physiol. Plant.,* 44, 417-421

Karunen, P. (1981) The role of neutral lipids in the physiology and ecology of subarctic *Dicranum elongatum. Can. J. Bot.,* 59, 1902-1909

Katagiri, T., Mizoguchi, T., & Shinozaki, K. (1996) Molecular cloning of a cDNA encoding diacylglycerol kinase (DGK) *in Arabidopsis thalia*na. *Plant Mol Biol.,* 30, 647-653

Katagiri, T., Takahashi, S., & Shinozaki, K. (2001) Involvement of a novel *Arabidopsis* phospholipase D, AtPLD□, in dehydration-inducible accumulation of phosphatidic acid in stress signalling. *Plant J.,* 26, 595-605

Kearns, E.V., Hugly, S., & Sommerville, C. (1991) The role of cytochrome b_5 in Δ^{12} desaturation of oleic acid by microsomes of safflowcr (*Carthamus tinctorius L.*). *Arch. Biochem. Biophys.,* 284, 431-436

Kohn, G., Vierengel, A., Vandekerkhove, O., & Hartmann, E. (1987a) 9-octadecen-6-ynoic acid from *Riccia fluitans. Phytochemistry,* 26, 2101.2102

Kohn, G., Demmerle, S., Vandekerkhove, O., Hartmann, E., & Beutelman, P. (1987b) Distribution and chemotaxonomic significance of acetylenic fatty acids in mosses of the dicranales. *Phytochemistry,* 26, 2271-2275

Kohn, G, Vandekerkhove, O., Hartmann, E., & Beutelmann, P. (1987c) Acetylenic fatty acids in the ricciaceae (hepaticae). *Phytochemistry,* 27, 1049-1051

Kolomiets, M.V., Hannapel, D.J., Chen, H., Tymeson, M., & Gladon, R.J. (2001) Lipoxygenase is involved in the control of potato tuber development. *Plant Cell,* 13, 613-626

Lamparter, T., Esch, H., Cove, D., Hughes, J., & Hartmann, E. (1996) Aphototropic mutants of the moss *Ceratodon purpureus* with spectrally normal and with spectrally dysfunctional phytochrome. *Plant Cell Environ.,* 19, 560-568

Lamparter, T., Esch, H., Cove, D., & Hartmann, E. (1997) Phytochrome control of phototropism and chlorophyll accumulation in the apical cells of protonemal filaments of wildtype and an aphototropic mutant of the moss *Ceratodon purpureus. Plant Cell Physiol.,* 38, 51-58

Lin, C. (2002) Blue light receptors and signal transduction. *Plant Cell, supplement* 2002, S207-S225

Malkönen. E. (1974) Annual primary production and nutrient cycle in some scots pine stands. *Commun. Inst. For. Fenn.,* 84, 1-87

Meijer, H.J.G., Berrie, C.P., Iurisci, C., Divecha, N., Musgrave, A., & Munnik, T. (2001) Identification of a new polyphosphoinositide in plants, phosphatidylinositol 5-monophosphate (Ptdins5P), and its accumulation upon osmotic stress. *Biochem. J.,* 360, 491-498

Meijer, H.J.G., Divecha, N., van der Ende, H., Musgrave, A., & Munnik, T. (1999) Hyperosmotic stress induces rapid synthesis of phosphatidyl-d-inositol 3,5-bisphosphate in plant cells. *Planta,* 208, 294-298

Meske, V., & Hartmann, E. (1995) Reorganization of microfilaments in protonemal tip cells of the moss *Ceratodon purpureus* during the phototropic response. *Protoplasma,* 188, 59-69

Meske, V., Ruppert, V., & Hartmann, E. (1996) Structural basis of the red light induced repolarization of tip growth in caulonema cells of *Ceratodon purpureus. Protoplasma,* 192, 189-198

Mikami, K., Katagiri, T., Iuchi, S., Yamaguchi-Shinozaki, K., & Shinozaki, K. (1998) A gene encoding phosphatidylinositol-4-phosphate 5-kinase is induced by water stress and abscisic acid in *Arabidopsis thaliana. Plant J.,* 15, 563-568

Millar, A.A., Wrscher, M., & Kunst, L. (1998) Accumulation of very-long-chain fatty acids in membrane glycerolipids is associated with dramatic alterations in plant morphology. Plant Cell, *11, 1889-1902*

Müller-Röber, B., & Pical, C. (2002) Inositol phospholipid metabolism in *Arabidopsis*. Characterized and putative isoforms of inositol phospholipid kinase and phosphoinositide-specific phospholipase C. *Plant Physiol.,* 130, 22-46

Munnik, T., Irvine, R.F., & Musgrave, A. (1989) Phospholipid signalling in plants. *Biochim. Biophys. Acta* ,1389, 222-272

Munnik, T. (2001) Phosphatidic acid: an emerging plant lipid second messenger. *Trends Plant Sci.,* 6, 227-33

Murata, N., & Tasaka, Y. (1997) Glycerol-3-phosphate acyltransferase in plants. *Biochim. Biophys. Acta* , 1348, 10-16

Nishizuka, Y. (1984) Turnover of inositol phospholipids and signal transduction. *Science,* 225, 1365-1370

Ohlrogge, J.B., & Jaworski, J.G. (1997) Regulation of fatty acid synthesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.,* 48, 109-136

Ohlrogge, J.B., & Jaworski, J.G. (1993) *De novo* fatty acid biosynthesis. In T.S. Moore, Jr.. (Ed.), *Lipid Metabolism in Plants* (pp3-32). Boca Raton, FL: CRC Press.

Otterhag, L., Sommarin, M., & Pical, C. (2001) N-terminal EF-hand-like domain is required for phosphoinositide-specific phospholipase C activity in *Arabidopsis thaliana. FEBS Lett., 497*, 165-170.

Pohl, P., & Wagner, H. (1972a) Fettsäuren im Pflanzen- und Tierreich (eine Übersicht) I: Gesättigte und cis-ungesättigte Fettsäuren. *Fette Seifen Anstrichmittel* 74, 424-435

Pohl, P., & Wagner, H. (1972b) Fettsäuren im Pflanzen- und Tierreich (eine Übersicht) II: Trans-ungesättigte, Alkin-, Hydroxy-, Epoxy-, Oxo-, Cyclopropan- und Cyclopropen-Fettsäuren. Fette Seifen Anstrichmittel *74, 541-550*

Pringsheim, E.G., & Pringsheim, O. (1935) Physiologische Studien an Moosen. Die Zuchtung von Laubmoosprotonemen im Dunkeln. *Jahrb. Wiss. Bot.,* 82, 312-332

Rebecchi, M.J., & Pentyala, S.N. (2000) Structure, function, and control of phosphoinositide-specific phospholipase C. *Physiol. Rev.,* 80, 1291-1335

Reddy, A.S., Nuccio, M.L., Gross, L.M., & Thomas, T.L. (1993) Isolation of a Δ^6-desaturase gene from the cyanobacterium *Synechocystis* sp. strain PCC 6803 by gain-of-function expression in *Anabaena* sp. strain PCC 7120. *Plant Mol Biol.*, 27, 293-300

Rhee, S.G. (2001) Regulation of phosphoinositide-specific phospholipase C. *Annu. Rev. Biochem.,* 70, 281-312

Samuelsson, B. (1983) Leucotrienes: mediators of immediate hypersensitivity and inflammation. *Science,* 220, 568-575

Sayanova, O., Smith, M.A., Lapinskas, P., Stobart, A.K., Dobson, G., Christie, W.W., Shewry, P.R., & Napier, J.A. (1997) Expression of a borage desaturase cDNA containing an N-terminal cytochrome b_5 domain results in the accumulatin of high levels of Δ6-desaturated fatty acids in transgenic tobacco. *Proc. Natl. Acad. Sci. USA,* 94, 4211-4216

Schaefer, D.G. (2002) A new moss genetics: Targeted mutagenesis in *Physcomitrella patens*. *Annu. Rev. Plant Biol.,* 53, 477-501

Schaefer, D.G., & Zryd, J.P. (1997) Efficient gene targeting in the moss P*hyscomitrella patens*. *Plant J.,* 11, 1195-1206

Schaefer, D.G., Zryd, J.P., Knight, C.D., & Cove, D.J. (1991) Stable transformation of the moss *Physcomitrella patens*. *Mol. Gen. Genet.,* 226, 418-424

Schmidt, H., & Heinz, E. (1990) Involvement of ferredoxin in desaturation of lipid-bound oleate in chloroplasts. *Plant Physiol.,* 94, 214-220

Shiran, D., Khozin, I., Heimer, Y.M., & Cohen, Z. (1996) Biosynthesis of eicosapentaenoic acid in the microalga *Porphyridium cruentum.* I. The use of externally supplied fatty acids. *Lipids,* 31, 1277-1282

Smith, H. (2000) Phytochromes and light signal perception by plants –an emerging synthesis. *Nature,* 407, 585-591

Smith, M.A., Cross, A.R., Jones, O.T.G., Griffiths, W.T, Stymne, S., & Stobart, K. (1990) Electron-transport components of the 1-acyl-2-oleoyl-*sn*-glycero-3-phosphocholine Δ^{12} desaturase (Δ^{12} desaturase) in microsomal preparations from developing safflower (*Carthamus tinctorius* L) cotyledons. *Biochem. J.*, 272, 23-29

Sopory, S.K., Sanan Neeti and Oelmüller. R (1999) Light signal transduction and Gene expression In G.S. Singal, G. Renger, S.K. Sopory, K.-D. Irrgang, & Govindjee (Eds.), *Concepts in Photobiology: Photosynthesis and Photomorphogenesis* (pp.897-927). New Delhi, India: Narosa Publishing House.

Stelmach, B.A., Müller, A., Hennig, P., Laudert, D., Andert, L., & Weiler, E.W. (1998) Quantitation of the octadecanoid 12-oxo-phytodienoic acid, a signalling compound in plant mechanotransduction. *Phytochemistry,* 47, 539-546

Stintzi, A. , Weber, H., Reymond, P., Browes, J., & Farmer, E.E. (2001) Plant defense in the absence of jasmonic acid: The role of cyclopentenones. Proc. Natl. Acad. Sci. USA, *98, 12837-12842*

Streb, H., Irvine, R.F., Berridge, M.J., & Schulz, I. (1983) Release of Ca^{2+} from non-mitochondrial intracellular store in pancreatic acinar cells by inositol-1,4,5-trisphosphate. *Nature,* 306, 67-69

Turner, J.T., Ellis, C., & Devoto, A. (2002) The jasmonate signal pathway. *Plant Cell, Supplement* 2002, S153-S164

Vandekerkhove, O., Euler, R., Kohn, G., & Hartmann, E. (1984) Influence of stress conditions on the fatty acid patterns of the moss *Leptobryum pyriforme. J. Hattori Bot. Lab.,* 56, 187-193

Vanhaesebroeck, B., Leevers, S.J., Ahmadi, K., Timms, J., Katso, R., Driscoll, P.C., Woscholski, R., Parker, P.J., & Waterfield, M.D. (2001) Synthesis and function of 3-phosphorylated inositol lipids. *Annu. Rev. Biochem.,* 70, 535-602

Voelker, D.R. (1997) Phosphatidylserine decarboxylase. *Biochim. Biophys. Acta,* 1348, 236-44

Voelker, T., & Kinney, A.J. (2001) Variations in the biosynthesis of seed-storage lipids. *Annu. Rev. Plant Physiol. Plant Mol. Biol.,* 52, 335-361

von Orlow, M. (2000) Isolation of cDNAs of the lipid metabolism from *Physcomitrella patens*. Master Thesis, Free University of Berlin, Department of Plant Physiology (Prof. E. Hartmann).

Wallis, J.G., & Browse, J. (2002) Mutants of Arabidopsis reveal many roles for membrane lipids. *Prog. Lipid Res.,* 41, 253-278

Wang, X. (2001) Plant phospholipases. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 52, 211-231

Zank, T.K., Zähringer, U., Beckmann, C., Pohnert, G., Boland, W., Holtorf, H., Reski, R, Lerchl, J., & Heinz, E. (2002) Cloning and functional characterization of an enzyme involved in the elongation of Δ6-polyunsaturated fatty acids from the moss *Physcomitrella patens. Plant J.,* 31, 255-268

Zhukova, N.V., & Aizdaicher, N.A. (1995) Fatty acid composition of 15 species of marine microalgae. *Phytochemistry,* 39, 351-356

PHYTOCHROME IN MOSSES

TILMAN LAMPARTER AND GERHARD BRÜCKER

Freie Universität Berlin, Pflanzenphysiologie, Königin Luise Str. 12-16, D-14195 Berlin, Germany

Abstract. Phytochromes are photoreceptors that are most sensitive in the red region of the visible spectrum. Their unique red-far-red photoreversibility has allowed early spectral detection and characterization in plants. In mosses, phytochrome controls spore germination, phototropism, polarotropism, side branch formation, chlorophyll accumulation, chloroplast redistribution and modulates gravitropism. Many of these responses take place in the protonemal tip cell and can easily be observed at the cellular level. Mutants with defects in phototropism have been isolated from *Physcomitrella patens* and *Ceratodon purpureus*. In class 1 mutants of *Ceratodon* the synthesis of phytochromobilin, which is the chromophore of phytochrome, is blocked. These mutants are defective in the conversion of heme to biliverdin, which is catalyzed by the enzyme heme oxygenase. Class 1 mutants have not only helped to study phytochrome action and biochemistry in *Ceratodon*. Their strong phenotype has also allowed to establish microinjection and gene replacement techniques for *Ceratodon*. Many vectorial cellular phytochrome responses are sensitive to polarized light. These findings are explained by dichroically oriented phytochrome molecules, which are most likely attached to the plasmamembrane. Aspects of this particular orientation will be discussed.

1. INTRODUCTION

Plants depend on light for photosynthesis. During development, many events are controlled by photoreceptor molecules that allow the plant to adapt to specific light conditions under various environmental conditions. These light-controlled processes, often summarized under the term photomorphogenesis, have been studied in great detail, especially during the last decade in the model plant *Arabidopsis thaliana*. Studies on algae, mosses and ferns have always contributed to our present knowledge of photoreceptor action and photomorphogenesis. Since mosses are less flexible in their development than seed plants, light induces

A.J. Wood et al. (ed.),
New Frontiers in Bryology: Physiology, Molecular Biology and Functional Genomics, 157–175.

subtler changes in mosses. In general, light-effects on protonemal and gametophytic cells are observed under a microscope. However, it is easier to study light effects at the cellular level with mosses, because many light effects can directly be observed in a single cell, the protonemal tip cell.

Three classes of photoreceptors are known in seed plants: cryptochromes, phototropins and phytochromes. The classes are distinguished by their spectral properties, protein sequence and choice of chromophore. Whereas phototropins have not been described for mosses, and database searches have failed to reveal a gene with clear phototropin homology, the other classes are represented by multiple genes. Cryptochromes are presented in chapter XIII, and phytochromes are addressed in this chapter.

Of the different plant photoreceptors, phytochromes were the first to be discovered. Phytochrome research dates back to the 1930s, when several scientists tested the spectral sensitivity of light-dependent processes such as seed germination and flower induction (Sage, 1992). It was discovered that such processes are induced by red light, and that long-wavelength red light, termed far-red light, was acting antagonistically. Initially these findings were explained by a proposed system consisting of two antagonistically acting photoreceptors. Later experiments, with alternating pulses of red and far-red light, showed that the far-red antagonism must be part of the photocycle of a single photoreceptor. It was discovered that, during the dark period which follows the red / far-red pulses, the effect on a process was determined by the nature of the last light pulse. In simple terms, the effect is switched on if the last light pulse is red, and off when the last light pulse is far-red. This indicated that the photoreceptor has two thermostable forms which are inter-converted by light: red switches the receptor to the active form, and far-red into the inactive form. It was also proposed that this unique feature, termed photoreversibility, could be the basis for spectral detection by the photoreceptor. Indeed, a spectral assay was developed to measure phytochrome in plant tissues and extracts (Butler et al., 1959), which in turn made purification of the chromoprotein possible (Siegelman and Firer, 1964). Subsequently, occurring earlier than the discovery of other plant photoreceptor genes by molecular genetic methods, the phytochrome gene was identified (Hershey et al., 1984).

There are many similarities between mosses and seed plants with regard to phytochrome responses. However, there are several remarkable differences. In seed plants, very low fluence, low fluence and high-irradiance phytochrome responses (VLFR, LFR and HIR, respectively) can be distinguished, but only the LFRs follow the classical pattern of phytochrome red/far-red reversibility. The VLFR and HIR are induced by far-red (Smith, 1994). It seems that in mosses all phytochrome responses follow the pattern of LFR, and there is no evidence that VLFRs or HIRs. VLFR and HIR are mediated by phytochrome A. It appears that phytochrome A is restricted to angiosperms and phytochrome A responses have only evolved within the evolution of angiosperms. Another difference between the groups relate to the vectorial phytochrome effects described below. We know

many examples of vectorial effects in mosses, but none in seed plants. Further study of phytochrome action in mosses will provide a deeper understanding of phytochrome evolution.

In the subsequent sections, we will first summarize the physiological responses of mosses that are mediated by phytochrome, present what is known of mutants with altered phytochrome responses, and finally discuss the molecular aspects of phytochrome physiology.

2. SPORE GERMINATION

The first evidence for the activity of a phytochrome in mosses derived from studies on spore germination. Spores of *Funaria hygrometrica* require light for germination. Germination is most sensitive to light in the red region of the spectrum and the red light effect can be reversed by far-red light (Bauer and Mohr, 1959). Red-far-red reversibility of germination has since been demonstrated for spores of *Ceratodon purpureus*, *Dicranum scoparium* (Valanne, 1966) and *Physcomitrella patens* (Schild, 1981).

3. PHOTOTROPISM AND POLAROTROPISM

Protonemata of many moss species exhibit a phytochrome-controlled phototropic response. The effect was first described in 1935 (Pringsheim and Pringsheim, 1935). Phototropism is a growth movement induced by a light stimulus and the growth direction is related to the direction of the incident light beam; the cells can either grow towards the light or away from it, depending on the species and physiological conditions. The role of phytochrome in vectorial responses such as phototropism appears to be restricted to cryptogams (less complex plants and ferns), for seed plants such responses are unknown (but see Dong et al., 1995, for chloroplast movements in *Vallisneria*). In angiosperm seedlings, phototropism is controlled by the blue light photoreceptor phototropin, red light has no vectorial effect but modulates the sensitivity of the phototropic response. Moss phototropism has been analyzed, in the most detail, for the species *Physcomitrella patens* (Jenkins and Cove, 1983b) and *Ceratodon purpureus* (Hartmann et al., 1983). At low light intensities, chloronematic filaments of *Physcomitrella* exhibit positive phototropism, that is they grow towards the light (Jenkins and Cove, 1983b). The response of caulonemal cells of *Physcomitrella,* however, is rather complex: there are always filaments that grow to and others that grow away from the light. At low light intensities, the majority grow away from the light. Intermediate light intensities induce a positive response in most filaments, and under strong light negative phototropism dominates (Brücker, 2003). *Ceratodon* filaments grow in a positively phototropic manner, irrespective of the light intensity (Hartmann et al., 1983; Esch et al., 1999). However, positive and negative curvature is obtained if the filaments are surrounded by water or agar

(Kern and Sack, 1999). An example of the phototropic response of a *Ceratodon* tip cell is shown in Figure 1a.

The phenomenon of polarotropism was initially described by Bünning and Etzold (1958). As the name suggests, polarotropism is a vectorial growth effect under the influence of polarized light. The response is comparable to the phototropic response, because in both cases the growth direction of the tip cell is determined by light. However, in the case of polarotropism, the growth direction is relative to the electrical (E) vector of polarized light and not to the directional vector of the light beam. Bünning and Etzold observed that protonemata of freshly germinated spores of *Funaria hygrometrica* aligned perpendicular to the E-vector of polarized light. The involvement of phytochrome was not investigated by these

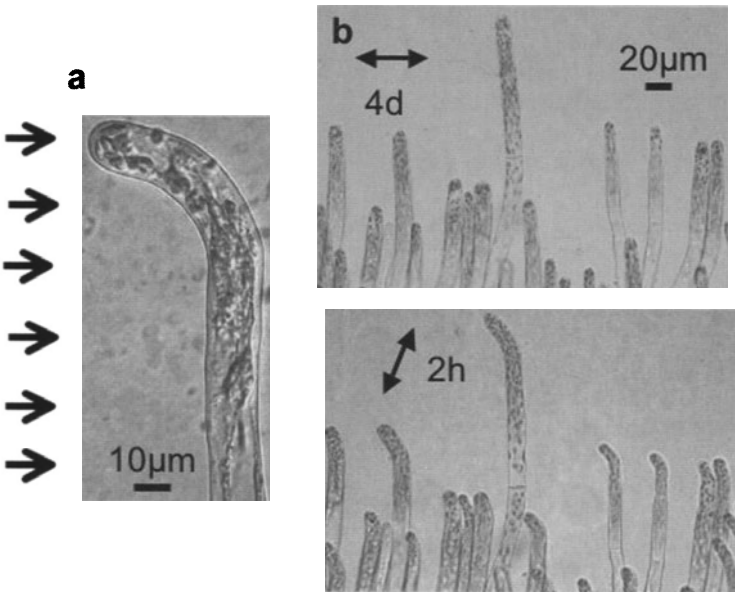

Figure 1 (a) Phototropism of a protonemal tip cell of *Ceratodon purpureus* (strain wt4). The cell was irradiated with monochromatic red light of 4 µmol m^{-2} s^{-1} from the left side (see arrows) for 2 h. Before this light treatment, the filaments were kept in unilateral low fluence white light (1 µmol m^{-2} s^{-1}) on a horizontally oriented agar dish for 4 days. This treatment leads to a parallel alignment of the filaments. In darkness, filaments may be aligned on vertically oriented agar dishes by their negative gravitropic growth; such a treatment is also required to align *ptr* mutants, which can not be aligned by light.

(b) Polarotropism of *Ceratodon* filaments. The filaments were grown for 4 days on a horizontally oriented agar dish with polarized red light (1 µmol m^{-2} s^{-1}) from above (top panel); the arrow indicates the orientation of the E-vector. Under such conditions, *Ceratodon* filaments grow perpendicular to the E-vector. Filaments that arise from the inoculation zone grow in two opposite directions, the panel shows the apical region of one growth direction. Note that the filaments are topped by a cover slip to avoid bending towards the light source (i.e. away from the growth medium). The lower panel shows the change of growth direction 2 h after turning the E-vector of polarized light. These experiments show that phytochrome molecules have a particular dichroic orientation within the cell.

authors, but has later been demonstrated as a factor for many other moss species such as *Physcomitrium turbinatum* (Nebel, 1969) *Physcomitrella* (Jenkins and Cove, 1983b) and *Ceratodon* (Esch et al., 1999). Polarized light certainly does not play an important role in the natural environment, thus polarotropism might be regarded as special type of phototropism that occurs under laboratory conditions and a response that may relate to a particular orientation of the phytochrome molecules within the cell (see below). Figure 1b shows an example of polarotropism of *Ceratodon* filaments.

4. PHYTOCHROME AND GRAVITROPISM

Moss filaments grow in a negatively gravitropic fashion (Jenkins et al., 1986; Schwuchow et al., 1995; Khorkavtsiv and Kardash, 2002; Schwuchow et al., 2002), i.e. upwards, when kept in darkness. In the light, phototropism dominates, masking the gravitropic response. However, the analyses of *Ceratodon* "class 2" mutants, which are defective in the phototropic response (see below), has facilitated the dissection of the effect of light on gravitropic growth from that of phototropic growth. Class 2 mutants grow in random directions under unilateral red light, indicating that red light inhibits gravitropism. Red/far-red experiments and analyses of class 1 mutants also confirmed the involvement of phytochrome (Lamparter et al., 1996) in gravitropism. Photo- and gravitropism studies, performed under variable *g* forces, demonstrate that both processes compete at low light intensities (Kern and Sack, 1999). The gravitropic response of *Ceratodon* is also altered by blue light. Since blue light induces only a very weak phototropic response in *Ceratodon*, the action of blue light on gravitropism is easier to analyze. It appears that blue light can invert the gravitropic response: under continuous irradiation with blue light the filaments grew downwards. This effect was independent of phytochrome (Lamparter et al., 1998a). There is an interesting parallel with the effect of light on gravitropism of seed plants in that gravitropism in *Arabidopsis* seedling shoots is randomized by light, an effect mediated both by phytochrome A and B (Poppe et al., 1996). The ecological significance of these effects is clear: in darkness, the most promising growth direction to find light is upwards, once the light is reached upward growth is no longer necessary.

5. CHLOROPLAST ORIENTATION

Light control of the intracellular distribution of chloroplasts is found in most plant cells and moss protonemal and gametophyte cells are no exception. High light intensities induce an avoidance response within cells that is designed to protect the chloroplast(s) from photodamage. Under low light intensities, however, chloroplasts move into positions that allow maximum photosynthetic light capture. In general, chloroplast movements are controlled by blue light photoreceptors, such as phototropin 1 and 2 as described for *Arabidopsis* (Kagawa

and Wada, 2002), but in cryptogam species, phytochrome has also been implicated. Of the two mosses that have been studied so far, both differ in this respect: in *Physcomitrella*, phytochrome does participate (Kadota et al., 2000), whereas in *Ceratodon*, it does not (Kagawa et al., 1997). As for the phototropic response, chloroplast redistribution is also a vectoral light response of individual cells, implicating dichroically oriented photoreceptor molecules. Classical experiments with polarized light performed with the green alga *Mougeotia* have led to the conclusion that active phytochrome molecules are located in the plasmamembrane (Haupt, 1960; Haupt, 1970).

6. CHLOROPHYLL ACCUMULATION & CHLOROPLAST FORMATION

In angiosperms, several steps in the synthesis of chlorophyll are light regulated. One of the last steps, the conversion from protochlorophyllide to chlorophyll, is in itself light dependent. It is because of this that angiosperms do not produce chlorophyll in the dark (Porra, 1997). Mosses however, do seem to be able to synthesize chlorophyll in the dark, although this has not been studied in great detail. Nevertheless, the chlorophyll content of dark-adapted *Ceratodon* tip cells is much lower than that of light-grown cells. Chlorophyll synthesis in this plant is under the control of phytochrome and a separate blue light photoreceptor (Lamparter et al., 1997). In addition, the shape of chloroplasts seems also to be controlled by light: plastids of dark-adapted tip cells are smaller than those of cells grown in red light.

7. FORMATION OF SIDE BRANCHES

During protonemal growth, basal cells can divide laterally and give rise to new tip cells. This side branch formation is controlled by light as shown for *Funaria* (Demkiv and Fedyk, 1977) *Physcomitrella* (Cove and Ashton, 1988) and *Ceratodon* (Kagawa et al., 1997). In the latter species the role for phytochrome was demonstrated by red/far-red treatments and mutant analyses. Polarized light had an interesting effect on phytochrome-induced branch formation in that both the position of the outgrowth and the branching frequency were dependent on the plane of vibration.

8. PROTOPLAST REGENERATION

Protoplast isolation and regeneration are useful tools in moss transformation. During regeneration, the round protoplasts form new cell walls and grow out at one particular side to form a new tip cell that subsequently forms the protonemal filament. The effect of light on *Ceratodon* protoplasts has been studied in detail in order to understand the formation of cell polarity (Cove et al., 1996). Both the regeneration rate and the position of outgrowth are light dependent. Of interest is the observation that the direction of outgrowth appears to be controlled almost

equally by both red and blue light. This implies that control is mediated by both phytochrome and a blue light photoreceptor. This contrasts with the phototropic response of filaments, where blue light induces only a subtle change of growth direction. Protoplasts do not respond to gravity, therefore they might respond more sensitive to gradients imposed by unidirectional light. This could explain why protoplasts respond stronger to blue light than protonemal tip cells.

9. PHOTOTROPISM MUTANTS

Phototropism mutants, designated *ptr*, have been isolated from both *Physcomitrella* (Cove et al., 1978) and *Ceratodon* (Lamparter et al., 1996). In the *Physcomitrella ptr* mutants, both phototropism and polarotropism are equally affected (Jenkins and Cove, 1983a), whereas other phytochrome responses are regulated normally (Cove and Lamparter, 1998). Based on genetic crosses, it was found that at least 3 different complementation groups contribute to the phototropic response. In *Ceratodon* two sets of mutants that have different phenotypes in continuous red light were isolated. Class 1 mutants grow upwards (negatively gravitropic) and produce only very low levels of chlorophyll, whereas class 2 mutants grow in random directions, and produce normal, wild-type levels of chlorophyll (see Figure 2). Protoplast fusion experiments demonstrate that representatives of either class complement each other (Lamparter et al., 1998b). It is not known, however, whether each mutant class contains several complementation groups. The class 2 mutants seem to be specifically affected in

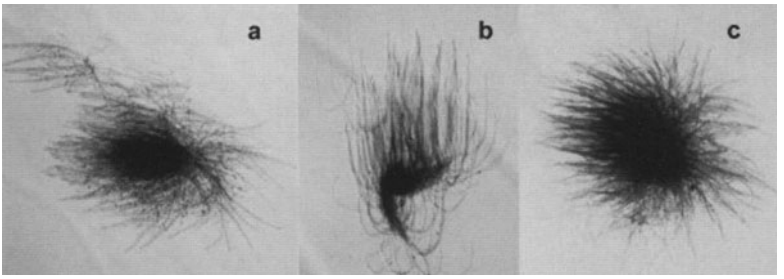

Figure 2 Phenotype of *Ceratodon* wild type and mutants under continuous red light. After inoculation, the filaments were grown on vertically oriented agar dishes for 5 days under unilateral red light (660 nm, 4 µmol m^{-2} s^{-1}) from the left. (a) wild type wt3 (b) class 1 mutant *ptr1* (c) class 2 mutant *ptr103*.

their phototropic response whilst other phytochrome controlled processes are regulated normally. In contrast, in class 1 mutants all phytochrome responses are lost. The nature of the lesion that confers the class 2 phenotype is as yet unclear. Class 1 mutants are defective in the biosynthesis of the phytochrome chromophore. All class 1 mutants can be rescued by either biliverdin (BV) or phycocyanobilin (PCB) (Esch and Lamparter, 1998). Biliverdin is a precursor in

the biosynthesis of phytochromobilin (PΦB), the natural phytochrome chromophore, and phycocyanobilin can functionally replace phytochromobilin (see Figure 3 for chromophore synthesis). These results suggest that the class 1 mutants are defective at the point of biliverdin formation. The relevant enzyme, heme oxygenase (HO), has been characterized in animals where it plays an important role in heme degradation and signaling processes. The first plant HO was found in *Arabidopsis* by mutant analyses (Muramoto et al., 1999).

Before the *Ceratodon* HO-gene was sequenced, HOs from rat and *Arabidopsis* were tested as to whether or not they could rescue the aphototropic phenotype of class 1 mutants. For this purpose, the respective genes were cloned into an expression vector and placed directly into mutant tip cells by means of microinjection (Brücker et al., 2000). This method had been previously established for the microinjection of low and high molecular dye components, phycocyanobilin, and GFP expression constructs into mutant and wild type cells. Injection of the HO-plasmids into *ptr116* cells resulted in a clear rescue of the mutant phenotype. The filaments formed during the days subsequent to the insertion grew towards unilateral red light and produced wild-type levels of chlorophyll. Microinjection is a powerful tool for the analysis of functional aspects of gene and metabolite activity at the cellular level, in the present case it allowed for a rapid assessment (within one day) of the effect of heme oxygenase expression on phytochrome mutant phenotypes (Brücker et al., 2000).

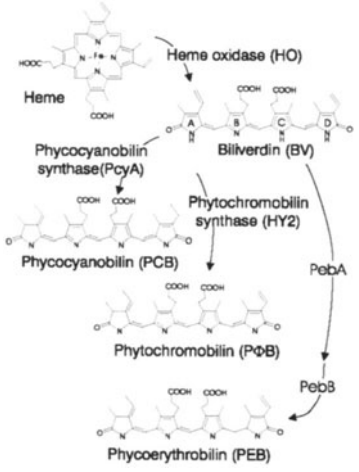

Figure 3. Last steps of bilin synthesis. Phytochromobilin is the chromophore of seed plants and moss phytochrome, whereas phycocyanobilin is used as the phytochrome chromophore in green algae and cyanobacteria. Both bilins can replace each other functionally. Phycoerythrobilin assembles with apo-phytochrome *in vitro* and *in vivo*. The PEB-adduct is highly fluorescent, but does not undergo photoconversion. In plants, chromophore synthesis takes place in the plastid. The cyanobacterial pathways leading to PCB and PEB are also shown. Both bilins are used as chromophores of phycobiliproteins which serve as photosynthetic antennae in cyanobacteria and red algae.

The full-length rat HO is membrane bound, but the construct used for the injection studies encodes a truncated, soluble protein that is expected to be located in the cytosol. The *Arabidopsis* HO contains a plastid target sequence and is transported into plastids, together with the other enzymes of PΦB-synthesis. The fact that both HOs rescue the mutant phenotype demonstrates that enzymatic activity in the cytosole is sufficient to bridge the gap between heme and PΦB.

These results show indirectly that heme and biliverdin can pass through the plastid membrane in both directions. However, in filaments that expressed the rat HO, phytochrome effects are weaker than in wild-type filaments or the filaments expressing *Arabidopsis* HO. In the former case, phototropic curvature was smaller and chlorophyll synthesis started later (after ca. 2 cell divisions).

Cloning and sequencing of a *Ceratodon* HO gene was achieved by using a PCR based cloning procedure with degenerate primers (F. Mittmann, unpublished). As expected, the deduced protein sequence contains a putative plastid-targeting domain. GFP fusion constructs indirectly confirm that the putative targeting domain does direct plastid localization. When moss vectors that encode for a HO-GFP fusion protein were injected into *Ceratodon* cells, GFP fluorescence was found exclusively in the plastid. In contrast, the protein was detected in the cytosol when the putative plastid targeting domain was deleted. Both the full length *Ceratodon* HO and the truncated enzyme rescued the mutant phenotype in a way similar to that described for the rat and *Arabidopsis* HOs. Both *Ceratoton* constructs in transgenic plants generate phototropic filaments with elevated chlorophyll levels. Curvature and chlorophyll synthesis in plants containing the cytosolic enzyme were lower than those expressing the plastidic version (Brücker, 2003) (see also Figure 4). Four to six days after transformation, the tip cell that originated from an injected filament lost its positive phototropic response, probably as the result of plasmid dilution. These cells did however produce chlorophyll for one more day.

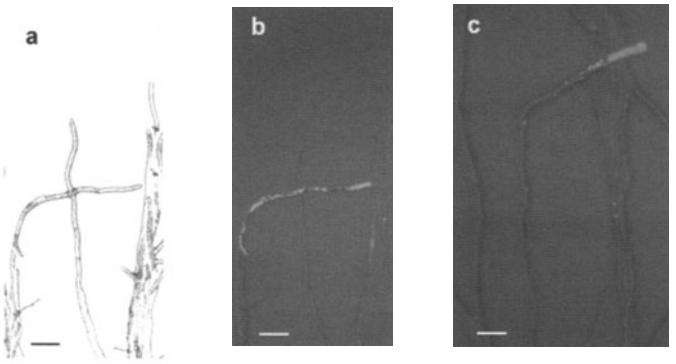

Figure 4. Rescue of phototropism and chlorophyll synthesis in filaments of the *Ceratodon* class 1 mutant *ptr115* by injection of expression constructs. (a) transmission and (b,c) chlorophyll autofluorescence. (a,b) An expression construct of the full-length HO was injected into one of the tip cells. Thereafter, the filaments were irradiated for 48 h (a,b) or 24 h (c) with red light (661 nm, 4 µmol m^{-2} s^{-1}) from the right. (c) a construct that encodes for truncated HO was injected into one of the cells.

The strong phenotype of *Ceratodon* class 1 mutants makes them ideal candidates for homologous recombination experiments in this species. In the plants, efficient homologous recombination has, so far, been restricted to *Physcomitrella patens* (see chapter IV,V and IX). Two different strategies

involving *Ceratodon* mutants have been employed: transformation of wild-type *Ceratodon* with a HO knockout construct to create a mutant phenotype or transformation of a class 1 mutant with a HO wild-type gene to rescue a mutant phenotype. Homologous recombination should generate clearly recognizable cell lines, either aphototropic filaments with low chlorophyll content in the first case, or phototropic filaments with high chlorophyll content in the second. When the first strategy was employed, the frequency of colonies with the anticipated phenotype was unexpectedly high in that 37% of the colonies were aphototropic after protoplast regeneration. However, PCR analysis of the selected lines revealed no evidence of homologous recombination. The phenotype appears to result from epigenetic gene silencing as seen in DNA methylation or RNAi experiments. When the second strategy was employed, homologous recombination in *Ceratodon* was achieved for the first time. Approximately 0.7 % of the regenerated protoplasts exhibited the wild-type phenotype (see Figure 5). PCR analysis and Southern blots confirmed that in these lines the mutant HO has been replaced by homologous recombination (Brücker, 2003). Since HO itself serves as reporter, this approach circumvents selection by antibiotic resistance. As a result, the selected lines are enriched for single insertions and illegitimate recombinations are unlikely. The findings demonstrate that gene replacement and site directed mutagenesis are possible in *Ceratodon*, at least under conditions that allow selection for the desired recombination event.

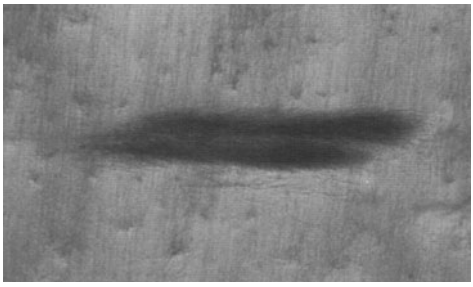

Figure 5. Rescue of *ptr116* (a class 1 mutant) after transformation with wild-type-like *Cerpu-HO* (the gene contains two silent mutations). Filaments were grown for 4 weeks on a vertically oriented agar plate with red light from the right. The *ptr116* filaments grow upwards, do not branch and have almost no chlorophyll. Filaments that originated from one rescued cell grow towards the light and contain high levels of chlorophyll. As a result of elevated photosynthesis and formation of side branches, rescued filaments increase their biomass rather rapidly. They can thus be easily separated from the non-rescued *ptr116* filaments. PCR and Southern blotting have shown that the rescue results from a homologous recombination event.

10. PHYTOCHROME GENES

The first cloned phytochrome gene of mosses, phytochrome 1 of *Ceratodon* (*CerpuPhy1*[1]), was exceptional (Thümmler et al., 1992). The deduced protein is homologous to known plant phytochromes in the N-terminal half, but divergent in the remaining C-terminal portion. The protein has a predicted molecular size of 145 kDa, significantly larger than common plant phytochromes, which range from 123 to 129 kDa. The C-terminal portion of CerpuPhy1 has similarities to protein kinases (see Figure 6), and serine/threonine kinase activity could be achieved *in vitro* for the truncated kinase domain in a fibroblast expression system (Thümmler et al., 1995). This finding brought new light to the discussion of plant phytochrome action, since the question as to whether or not plant phytochromes act as protein kinases was a controversial one. Another exceptional phytochrome was also found in the fern *Adiantum capillus veneris* (Nozue et al., 1998). This phytochrome, AdicaPhy3, also contains a kinase in its C-terminus but is divergent from that of CerpuPhy1. The entire C-terminal region of AdicaPhy3 has homology to phototropin. In addition, besides the kinase domain, this region contains two LOV domains that are known to bind flavin chromophores (see Figure 6). Thus, AdicaPhy3 binds three chromophores, a bilin in the N-terminus and two flavins in the C-terminal region (Christie et al., 1999; Nozue et al., 1998).

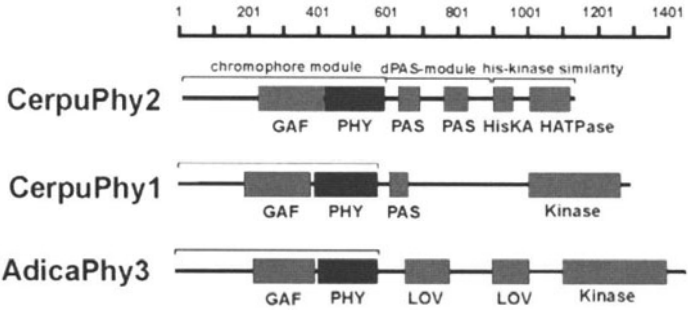

Figure 6. Domain structure of phytochromes. CerpuPhy2 (phytochrome 2 of *Ceratodon purpureus*) stands for "conventional" phytochromes that are known from seed plants. All other known moss phytochromes with the exception of the unconventional CerpuPhy1 also belong to this type. The domain structure of the unconventional phytochrome 3 from the fern *Adiantum capillus veneris* (AdicaPhy3) is also shown.

In *Arabidopsis*, phototropin controls phototropism, but also many other blue-light responses. Mutant analyses and silencing studies suggest that AdicaPhy3 controls

[1] The nomenclature and abbreviations used here may differ from the cited articles, because there is as yet no general rule for lower plant phytochromes. The abbreviations stand for the species name and contain a number for the phytochrome gene. The numbering indicates the chronological order of sequencing, but has no functional meaning. Italic abbreviations stand for the gene, the proteins are given in normal letters.

the red-light mediated phototropic response of the fern protonema (Kadota and Wada, 1999, Kawai et al., 2003).

The function of CerpuPhy1 is as yet unclear. It appears to be expressed at low levels, since its mRNA is difficult to detect (Pasentsis et al., 1998). In addition, antibodies derived against the N-terminal region failed to detect the protein but did recognize a smaller polypeptide (Lamparter et al., 1995). These data indicated the presence of another phytochrome gene in *Ceratodon*. Using the 5' part of *CerpuPhy1* as a probe, a second phytochrome was identified by Southern blots and finally cloned and sequenced (Hughes et al., 1996). This gene encodes for a "normal" phytochrome, which is homologous over its entire length to known seed plant orthologs. The first *Physcomitrella* phytochrome gene isolated also encodes a normal plant phytochrome (Kolukisaoglu et al., 1993). PCR and Southern blotting have revealed further phytochrome genes from both species, in total there are now three fully sequenced phytochrome genes in *Ceratodon* and four in *Physcomitrella* (Mittmann, 2003). Of three ESTs sequenced in the PEP EST program, two are identical with *PhypaPhy2* and *4*, but the third probably represents a fifth phytochrome gene. Four of the *Physcomitrella* phytochromes were knocked out (Mittmann, 2003) and the filaments tested for phototropism (Brücker, 2003). All knockout lines exhibit a growth pattern divergent from the wild type, indicating that all products participate in the phototropic response.

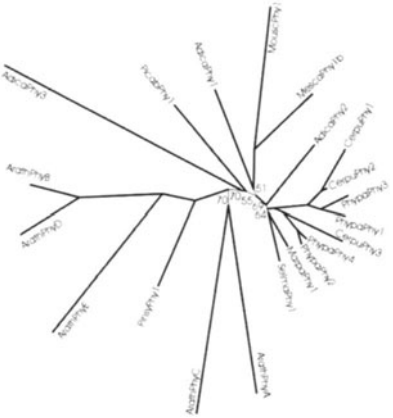

Figure 7. Phylogenetic tree of phytochromes from mosses and other plants. The tree was constructed from protein sequences of the N-terminal chromophore module (see Figure 6). Alignment of the proteins was performed with ClustalX (Thompson et al., 1997) using default parameters. The tree was constructed with the PHYLIP program package (Felsenstein, 2000) using the PROTDIST and FITCH algorithms. For bootstrapping, 100 datasets of sequences were generated, bootstrap values were usually above 90%. Branches with bootstrap values < 75% are drawn as thin lines, in those cases the % values are indicated. The following abbreviations were used: Adica: *Adiantum capillus veneris*, Arath: *Arabidopsis thaliana*, Cerpu: *Ceratodon purpureus*, Mesca: *Mesotaenium caldariorum*, Marpo: *Marchantia polymorpha*, Mousc: *Mougeotia scalaris*, Phypa: *Physcomitrella patens*, Picab: *Picea abies*, Pinys: *Pinus sylvestris*, Selma: *Selaginella martensii*

A phylogenetic tree, constructed from protein sequences, is given in Figure 7. According to this analysis, two groups of phytochromes can be distinguished in mosses, one group consisting of CerpuPhy1, CerpuPhy2, PhypaPhy1 and PhypaPhy3, the other CerpuPhy3, PhypaPhy2 and PhypaPhy4. It seems that *CerpuPhy1* and *CerpuPhy2* arose from a late gene duplication and subsequent domain re-arrangement to form *CerpuPhy1*. It is no surprise, therefore, that phytochrome genes that are similar to *CerpuPhy1* have not been found in other species. The *Physcomitrella* pairs *PhypaPhy1/3* and *PhypaPhy2/4* have also arisen from a duplication event after the divergence of the two groups of moss phytochrome genes. The simplest explanation for this arrangement is a duplication of the entire *Physcomitrella* genome. This assumption is consistent with the finding that the genome of *Ceratodon* is smaller (ca 250 Mbp) than that of *Physcomitrella* (Lamparter et al., 1998b).

11. PROTEIN STUDIES

There are few studies of moss phytochromes in regards to their properties as proteins. CerpuPhy2 has been expressed in a recombinant *S. cerevisiae* system and, as for other phytochromes, it autocatalytically attaches the chromophores PCB, PEB and PΦB and behaves as a homodimer (Zeidler et al., 1998). The PΦB adduct gives a difference spectrum identical to that of partially purified phytochrome from *Ceratodon* filaments (Lamparter et al., 1995). This implies that the natural chromophore is PΦB, as in seed plants. Spectral properties were reminescent of B-type angiosperm phytochromes (Zeidler et al., 1998; Sineshchekov et al., 2000). A polyclonal antibody, Apc1, has been produced against the N-terminal chromophore domain of CerpuPhy1 and used for various studies (Lamparter et al., 1995). As mentioned above, the CerpuPhy1 band was not detected on immunoblots, but a protein of smaller size was. The high sequence similarity between CerpuPhy1 and CerpuPhy2 implies that the band relates to CerpuPhy2. The immobilized antibody also precipitated photoreversible phytochrome from the crude extract of *Ceratodon* filaments. However, around 50% of the spectral activity was not precipitated (Lamparter et al., 1995). This remaining spectral activity could be related to CerpuPhy3.

The Apc1 antibody was used to study the concentration of phytochrome and its subcellular localization. It is known that spectrally detectable phytochrome in *Ceratodon* is light regulated, in dark adapted filaments, the concentration is higher than in filaments grown in the light (Lamparter et al., 1995). This light /dark pattern was confirmed by immunoblots. In aclass 1 mutant however, the content of apophytochrome was equally high in both dark- and light-grown filaments (Esch and Lamparter, 1998). This implies that the phytochrome chromophore is required for the regulation of phytochrome content. Most likely, CerpuPhy2 is slowly degraded in its far-red absorbing Pfr form (which is the active form of

phytochrome). This is reminiscent of angiosperm phytochrome A, which is also degraded in its Pfr form, however at a faster rate (Clough and Vierstra, 1997).

Soluble and pelletable fractions of *Ceratodon* extracts were tested for their cross-reactivity with Apc1. The majority of the phytochrome was found in the soluble fraction, but about 15% was detected in the pelletable membrane fraction (Esch and Lamparter, 1998). When the Apc1 antibody was used for immunocytological assays, phytochrome was distributed uniformly throughout the cytosol (T. Lamparter, unpublished).

The specific role(s) for each phytochrome protein is as yet unclear. The fact that none of the isolated aphototropic mutants of *Ceratodon* or *Physcomitrella* appear to be defective for a particular phytochrome suggests redundancy in the phototropic response. The knockout mutants of *Physcomitrella* are all altered in their phototropic response, but none were specifically aphototropic. Thus, none of the known single phytochromes is absolutely required for phototropism, and either two or more can substitute for the missing member of the family. The possibility that there is still an unidentified single phytochrome that could be responsible for the phototropic response of *Ceratodon* or *Physcomitrella* cannot be ruled out.

12. DICHROIC ORIENTATION

Phototropism, polarotropism, chloroplast reorientation and side branch initiation point to dichroically oriented phytochrome in the cell. As mentioned above, it is expected that phytochrome molecules that are located in close vicinity to the plasmamembrane mediate these responses. However, most reports concerning the intracellular localization of phytochrome conflict with that assumption.

Phytochromes are somewhat hydrophilic proteins that lack transmembrane domains. They are extracted largely as soluble molecules, and early imunological studies on intracellular distribution in seed plants showed that the majority is located in the cytosole (Pratt, 1994). For seed plants it was demonstrated that phytochromes can enter the nucleus. These movements were initially analyzed with fusion constructs containing either GUS (Sakamoto and Nagatani, 1996) or GFP (Kircher et al., 1999), and were later confirmed by classical antibody studies (Hisada et al., 2000). An *in vitro* membrane association of phytochrome has been studied for some decades. Several percent of extracted phytochromes are always found in the pelletable membrane fraction (Napier and Smith, 1987; Lamparter et al., 1992), as seen for extracts of *Ceratodon* (Podlowski, 1994; Esch and Lamparter, 1998). Most of the pelleted phytochrome becomes soluble in the presence of mild membrane-solubilizing detergents (Lamparter et al., 1992; Podlowski, 1994). This indicates that an association to other insoluble structures such as the cytoskeleton does not account for the pelletability of the phytochromes. It is still unclear whether or not the *in vitro* membrane association represents a phytochrome fraction that is bound to the plasmamembrane *in vivo*.

Molecular investigations may reveal the presence of membrane proteins that may account for this association.

As mentioned previously, the Apc1 antibody has been used for immunocytological detection of phtochrome proteins in *Ceratodon* tip cells; phytochrome was also detected in the cytosole in this case. Another labeling strategy was also used in *Ceratodon* cells. Phycoerythrobilin, which can be extracted from red algae, is used to replace the natural chromophore of phytochrome. Since PEB lacks the double bond between ring C and D (Figure 3), the adduct does not undergo photoconversion. Instead, the absorbed energy is converted at a high efficiency into fluorescence. This labeling technique has been used in studies of *Arabidopsis* seedlings (Murphy and Lagarias, 1997). Since the natural chromophore competes with PEB, chromophore-deficient mutants give a much stronger signal than wild-type plants. The same was found when *Ceratodon* class 1 mutants and wild-type cells were compared (T. Lamparter, unpublished). As shown in Figure 8, phytochrome was uniformly distributed in the cytosol.

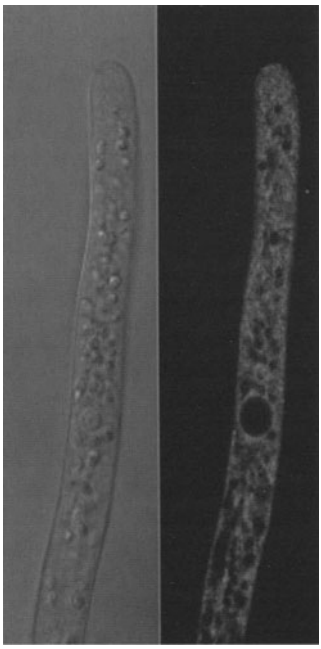

Figure 8. Intracellular phytochrome distribution in a tip cell of the class 1 mutant ptr116 after PEB labeling. Left: transmission image, right: fluorescence of the PEB-phytochrome adduct. Before fluorescence assay, cells are kept for 24 h on agar-medium containing 6 μM PEB at 20°C.

With fluorescence correlation spectroscopy (FCS) it is possible to measure the mobility of fluorescing particles at particular position within the cell. In PEB-

fed *Ceratodon* tip cells the mobility of phytochrome at the periphery of the cell was almost 10 times lower that that seen in the remaining parts of the cell (Böse et al., 2003). This was strong evidence for a membrane associated phytochrome in *Ceratodon* tip cells and can be developed for use in other species with either PEB- or GFP-labeled phytochrome. The reason why other methods failed to detect phytochrome at the cell periphery *in vivo* might lie in the relatively high content of cytosolic phytochrome; within that background it would be impossible to recognize the small fraction of membrane associated molecules by conventional methods.

13. SIGNAL TRANSDUCTION

Moss protonemata offer the opportunity to monitor signal transduction at the cellular level. The putative second messenger Ca^{2+} was detected in tip cells of *Ceratodon* with chlorotetracycline, a membrane bound fluorescence reporter molecule. Dark grown filaments develop a strong tip-to base Ca^{2+} gradient. When red light was given from the side, a rapid shift in the Ca^{2+} gradient towards the side of prospective growth was observed (Hartmann and Weber, 1988; Meske and Hartmann, 1995). This suggests that Ca^{2+} acts as a second messenger in the phytochrome-mediated phototropic response. A phytochrome effect on membrane potential and on calcium channels has been described for *Phycomitrella* (Ermolayeva et al., 1996; Ermolayeva et al., 1997).

The phototropic response requires an intact actin cytoskeleton, but is independent of microtubules, as shown by inhibitor studies. The actin cytoskeleton is arranged in a characteristic manner around the growing zone of the tip cell. The rearrangement of the actin cytoskeleton, observed after a unilateral light stimulus, also precedes outgrowth towards the new growth direction (Meske et al., 1996). Thus, for a change of growth direction during phototropism in *Ceratodon* to occur, the following signal transduction sequence is proposed: phytochrome / Ca^{2+} / actin cytoskeleton / change of growth direction. Further details of the signal transduction cascade remain as yet unclear, but phosphatidyl-inositole metabolism and phospholipase C might play a role in phototropism and other phytochrome responses (Pfaffmann and Hartmann, 1988). For recent details on this particular lipid metabolism see chapter XI. From protoplast regeneration (Cove et al., 1996), microbeam experiments (Esch et al., 1999) and inhibitor studies (Hartmann and Weber, 1988) it is clear that the phototropic stimulus can be memorized under certain conditions. Thus, a gradient, which mirrors the direction of previous light ,must stay fixed in the tip cell for several hours.

14. CONCLUSIONS

Based on the molecular research accomplished during the last decade or so, many details concerning phytochrome action have been revealed. For example, several phytochrome-interacting proteins have been identified and shown to play a

role in light-controlled processes. However, the picture that emerges is complex and many questions still remain. Moss protonemata provide a model in which phytochrome action can be observed with ease at the cellular level. Molecular, biochemical and physiological moss studies have already contributed much to our present knowledge of phytochrome action, and should still help to unravel the intricacies of phytochrome function and signal transduction in plants.

15. REFERENCES

Bauer, L., & Mohr, H. (1959). Der Nachweis des reversiblen Hellrot-Dunkelrot Reaktionssystems bei Laubmoosen. Planta *54*, 68-73.

Böse, G., Schwille, P., & Lamparter, T. (2003). The mobility of phytochrome within protonemal tip cells of the moss *Ceratodon purpureus*, monitored by fluorescence correlation spectroscopy, *submitted*

Brücker, G. (2003) Molekular- und zellbiologische Untersuchungen zum Phototropismus in Moosen. PhD Thesis Freie Universität Berlin, Germany.

Brücker, G., Zeidler, M., Kohchi, T., Hartmann, E., & Lamparter, T. (2000). Microinjection of heme oxygenase genes rescues phytochrome-chromophore-deficient mutants of the moss *Ceratodon purpureus*. Planta *210*, 529-535.

Butler, W.L., Norris, K.H., Siegelman, H.W., & Hendricks, S.B. (1959). Detection, assay, and preliminary purification of the pigment controlling photoresponsive development of plants. Proc.Nat.Acad.Sci.USA *45*, 1703-1708.

Bünning, E., & Etzold, H. (1958). Über die Wirkung von polarisiertem Licht auf keimende Sporen von Pilzen, Moosen und Farnen. Ber.Dt.Bot.Ges. *71*, 304-306.

Christie, J.M., Salomon, M., Nozue, K., Wada, M., & Briggs, W.R. (1999). LOV (light, oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): binding sites for the chromophore flavin mononucleotide. Proc.Natl.Acad.Sci.U S A *96*, 8779-8783.

Clough, R.C., & Vierstra, R.D. (1997). Phytochrome degradation. Plant Cell Environ. *20*, 713-721.

Cove, D., Schild, A., Ashton, N.W., & Hartmann, E. (1978). Genetic and physiological studies of the effect of light on the development of the moss, *Physcomitrella patens*. Photochem. Photobiol. *27*, 249-254.

Cove, D.J., & Ashton, N.W. (1988). Growth regulation and development in *Physcomitrella patens*: an insight into growth regulation and development in bryophytes. Botanical Journal of the Linnean Society *98*, 247-254.

Cove, D.J., & Lamparter, T. (1998). Light and moss development. In Microbial responses to light and time. M.X. Caddick, S. Baumberg, D.A. Hodgson & M.K. Phillips-Jones, eds. (Cambridge: Cambridge University Press), pp. 125-141.

Cove, D.J., Quatrano, R.S., & Hartmann, E. (1996). The alignment of the axis of asymmetry in regenerating protoplasts of the moss, Ceratodon purpureus, is determined independently of axis polarity. Development *122*, 371-379.

Demkiv, O.T., & Fedyk, I.D. (1977). Polarity of cell permeability and its control by phytochrome. Biofizika *3*, 824-828.

Dong, X.J., Takagi, S., & Nagai, R. (1995). Regulation of the orientation movement of chloroplasts in epidermal cells of *Vallisneria*: Cooperation of phytochrome with photosynthetic pigment under low-fluence-rate light. Planta *197*, 257-263.

Ermolayeva, E., Hohmeyer, H., Johannes, E., & Sanders, D. (1996). Calcium-dependent membrane depolarisation activated by phytochrome in the moss *Physcomitrella patens*. Planta *199*, 352-358.

Ermolayeva, E., Sanders, D., & Johannes, E. (1997). Ionic mechanism and role of phytochrome-mediated membrane depolarisation in caulonemal side branch initial formation in the moss *Physcomitrella patens* . Planta *201*, 109-118.

Esch, H., Hartmann, E., Cove, D., Wada, M., & Lamparter, T. (1999). Phytochrome-controlled phototropism of protonemata of the moss *Ceratodon purpureus*: physiology of wild type and class 2 *ptr* mutants. Planta *209*, 290-298.

Esch, H., & Lamparter, T. (1998). Light regulation of phytochrome content in wild-type and aphototropic mutants of the moss *Ceratodon purpureus*. Photochem.Photobiol. *67*, 450-455.

Felsenstein, J. (2000) PHYLIP (Phylogeny Inference Package) Version 3.6. . University of Washington, Seattle, Department of Genetics, distributed by the author.

Hartmann, E., Klingenberg, B., & Bauer, L. (1983). Phytochrome mediated phototropism in protonemata of the moss *Ceratodon purpureus* BRID. Photochem.Photobiol. *38*, 599-603.

Hartmann, E., & Weber, M. (1988). Storage of the phytochrome-mediated phototropic stimulus of moss protonematal cells. Planta *175*, 39-49.

Haupt, W. (1960). Die Chloroplastendrehung bei Mougeotia. II: Die Induktion der Schwachlichtbewegung durch linear polarisiertes Licht. Planta *55*, 465-479.

Haupt, W. (1970). Localization of phytochrome in the cell. Physiol.Veg. *8*, 551-563.

Hershey, H.P., Colbert, J.T., Lissemore, J.L., Barker, R.F., & Quail, P.H. (1984). Molecular cloning of cDNA for *Avena* phytochrome. Proc.Nat.Acad.Sci.USA *81*, 2332-2336.

Hisada, A., Hanzawa, H., Weller, J.L., Nagatani, A., Reid, J.B., & Furuya, M. (2000). Light-induced nuclear translocation of endogenous pea phytochrome A visualized by immunocytochemical procedures. Plant Cell *2000*, 1063-1078.

Hughes, J., Lamparter, T., & Mittmann, F. (1996). Cerpu;PHY0;2, a "normal" phytochrome in *Ceratodon*. Plant Physiol. *112*, 446-446.

Jenkins, G.I., Courtice, G.R., & Cove, D.J. (1986). Gravitropic responses of wild-type and mutant strains of the moss *Physcomitrella patens*. Plant Cell Environ. *9*, 637-644.

Jenkins, G.I., & Cove, D.J. (1983a). Phototropism and polarotropism of primary chloronemata of the moss *Physcomitrella patens*: responses of mutant strains. Planta *159*, 432-438.

Jenkins, G.I., & Cove, D.J. (1983b). Phototropism and polarotropism of primary chloronemata of the moss *Physcomitrella patens*: responses of the wild-type. Planta *158*, 357-364.

Kadota, A., Sato, Y., & Wada, M. (2000). Intracellular chloroplast photorelocation in the moss *Physcomitrella patens* is mediated by phytochrome as well as by a blue-light receptor. Planta *2000*, 932-937.

Kadota, A., & Wada, M. (1999). Red light-aphototropic (Rap) mutants lack red light-induced chloroplast relocation movement in the fern *Adiantum capillus veneris*. Plant Cell Physiol. *40*, 238-247.

Kagawa, T., Lamparter, T., Hartmann, E., & Wada, M. (1997). Phytochrome-mediated branch formation in protonemata of the moss *Ceratodon purpureus*. J.Plant Res. *110*, 363-370.

Kagawa, T., & Wada, M. (2002). Blue light-induced chloroplast relocation. Plant Cell Physiol *43*, 367-371.

Kawai, H., Kanegae, T., Christensen, S., Kiyosue, T., Sato, Y., Imaizumi, T., Kadota, A., & Wada, M. (2003). Responses of ferns to red light are mediated by an unconventional photoreceptor. Nature 421, 287-290.

Kern, V.D., & Sack, F.D. (1999). Irradiance-dependent regulation of gravitropism by red light in protonemata. Planta *1999*, 299-307.

Khorkavtsiv, O.Y., & Kardash, O.R. (2002). Gravity-dependent reactions of the moss *Pohlia nutans* protonemata. Adv.Space.Res. *27*, 989-993.

Kircher, S., Kozma-Bognar, L., Kim, L., Adam, E., Harter, K., Schäfer, E., & Nagy, F. (1999). Light quality-dependent nuclear import of the plant photoreceptors phytochrome A & B. Plant Cell *11*, 1445-1456.

Kolukisaoglu, H.U., Braun, B., Martin, W.F., & Schneider-Poetsch, H.A.W. (1993). Mosses do express conventional, distantly B-type-related phytochromes. Phytochrome of *Physcomitrella patens* (Hedw.). FEBS Lett. *334*, 95-100.

Lamparter, T., Brücker, G., Esch, H., Hughes, J., Meister, A., & Hartmann, E. (1998b). Somatic hybridisation with aphototropic mutants of the moss *Ceratodon purpureus*: genome size, phytochrome photoreversibility, tip-cell phototropism and chlorophyll regulation. J.Plant Physiol. *153*, 394-400.

Lamparter, T., Esch, H., Cove, D., & Hartmann, E. (1997). Phytochrome control of phototropism & chlorophyll accumulation in the apical cells of protonemal filaments of wildtype and an aphototropic mutant of the moss *Ceratodon purpureus*. Plant Cell Physiol. *38*, 51-58.

Lamparter, T., Esch, H., Cove, D., Hughes, J., & Hartmann, E. (1996). Aphototrophic mutants of the moss *Ceratodon purpureus* with spectrally normal and with spectrally dysfunctional phytochrome. Plant Cell Environ. *19*, 560-568.

Lamparter, T., Hughes, J., & Hartmann, E. (1998a). Blue light- and genetically-reversed gravitropic response in protonemata of the moss *Ceratodon purpureus*. Planta *206*, 95-102.

Lamparter, T., Lutterbuese, P., Schneider-Poetsch, H.A.W., & Hertel, R. (1992). A study of membrane-associated phytochrome: Hydrophobicity test and native size determination. Photochem.Photobiol. *56*, 697-707.

Lamparter, T., Podlowski, S., Mittmann, F., Schneider-Poetsch, H., Hartmann, E., & Hughes, J. (1995). Phytochrome from protonemal tissue of the moss *Ceratodon purpureus*. J.Plant Physiol. *147*, 426-434.

Meske, V., & Hartmann, E. (1995). Reorganisation of microfilaments in protonemal tip cells of the moss *Ceratodon purpureus* during the phototropic response. Protoplasma *188*, 59-69.

Meske, V., Rupert, V., & Hartmann, E. (1996). Structural basis for the red light induced repolarisation of tip growth in caulonemal cells of *Ceratodon purpureus*. Protoplasma *192*, 189-198.

Mittmann, F. (2003) Molekularbiologische Untersuchungen zum Phytochromsystem der Moose *Physcomitrella patens* und *Ceratodon purpureus*. PhD Thesis Freie Universität Berlin, Germany.

Muramoto, T., Kohchi, T., Yokota, A., Hwang, I., & Goodman, H.M. (1999). The Arabidopsis photomorphogenic mutant hy1 is deficient in phytochrome chromophore biosynthesis as a result of a mutation in a plastid heme oxygenase. Plant Cell *11*, 335-348.

Murphy, J.T., & Lagarias, J.C. (1997). The phytofluors: a new class of fluorescent protein probes. Curr.Biol. *7*, 870-876.

Napier, R.M., & Smith, H. (1987). Photoreversible association of phytochrome with membranes. II Reciprocity tests and a model for the binding reaction. Plant Cell Environ. *10*, 391-396.

Nebel, B.J. (1969). Responses of moss protonemata to red and far-red polarized light: evidence for disc-shaped phytochrome photoreceptors. Planta *87*, 170-179.

Nozue, K., Kanegae, T., Imaizumi, T., Fukuda, S., Okamoto, H., Yeh, K.C., Lagarias, J.C., & Wada, M. (1998). A phytochrome from the fern Adiantum with features of the putative photoreceptor NPH1. Proc.Natl.Acad.Sci.U S A *95*, 15826-15830.

Pasentsis, K., Paulo, N., Algarra, P., Dittrich, P., & Thümmler, F. (1998). Characterization and expression of the phytochrome gene family in the moss *Ceratodon purpureus*. Plant J. *13*, 51-61.

Pfaffmann, H., & Hartmann, E. (1988). Analysis of phospholipids and assay for phospholipases from moss tissue. In Methods in Bryology. J.M. Glime, ed. (Nichinan: Hattori Bot. Lab.), pp. 177-180.

Podlowski, S. (1994) Biochemische und immunologische Untersuchungen zum Photorezeptor Phytochrom aus dem Laubmoos *Ceratodon purpureus*. Diploma thesis, Freie Universität Berlin, Germany.

Poppe, C., Hangarter, R.P., Sharrock, R.A., Nagy, F., & Schäfer, E. (1996). The light-induced reduction of the gravitropic growth- orientation of seedlings of *Arabidopsis thaliana* (L.) Heynh. is a photomorphogenic response mediated synergistically by the far-red- absorbing forms of phytochromes A and B. Planta *199*, 511-514.

Porra, R.J. (1997). Recent progress in porphyrin and chlorophyll biosynthesis. Photochem. Photobiol. *65*, 492-516.

Pratt, L.H. (1994). Distribution and localization of phytochrome within the plant. In Photomorphogenesis in plants. R.E. Kendrick & G.H.M. Kronenberg, eds. (Dortrecht: Kluwer), pp. 163-186.

Pringsheim, E.G., & Pringsheim, G. (1935). Physiologische Studien an Moosen. 3. Die Züchtung von Laubmoosprotonemen im Dunkeln. Jahrbücher für wissensch.Bot. *82*, 312-332.

Sage, L.C. (1992). Pigment of the Imagination - A History of Phytochrome Research (Academic Press).

Sakamoto, K., & Nagatani, A. (1996). Nuclear localization activity of phytochrome B. Plant J. *10*, 859-868.

Schild, A. (1981) Untersuchungen zur Sporenkeimung und Protonemaentwicklung bei dem Laubmoos *Physcomitrella patens*. PhD thesis, Universität Mainz, Germany.

Schwuchow, J.M., Kern, V.D., White, N.J., & Sack, F.D. (2002). Conservation of the plastid sedimentation zone in all moss genera with known gravitropic protonemata. J.Plant Growth Regul. *21*, 146-155.

Schwuchow, J.M., Kim, D., & Sack, F.D. (1995). Caulonemal gravitropism and amyloplast sedimentation in the moss *Funaria*. Can.J Bot. *73*, 1029-1035.

Siegelman, H.W., & Firer, E.M. (1964). Purification of phytochrome from oat seedlings. Biochemistry *3*, 418-423.

Sineshchekov, V., Koppel, L., Hughes, J., Lamparter, T., & Zeidler, M. (2000). Recombinant phytochrome of the moss *Ceratodon purpureus* (CP2): fluorescence spectroscopy and photochemistry. J.Photochem.Photobiol.B *56*, 145-153.

Smith, H. (1994). Sensing the light environment: the functions of the phytochrome family. In Photomorphogenesis in plants. R.E. Kendrick & G.H.M. Kronenberg, eds. (Dortrecht: Kluwer), pp. 377-416.

Thompson, J.D., GibsonT.J., Plewniak, F., Jeanmougin, F., & Higgins, D.G. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. *24*, 4876-4882.

Thümmler, F., Dufner, M., Kreisl, P., & Dittrich, P. (1992). Molecular cloning of a novel phytochrome gene of the moss *Ceratodon purpureus* which encodes a putative light-regulated protein kinase. Plant Mol.Biol. *20*, 1003-1017.

Thümmler, F., Herbst, R., Algarra, P., & Ullrich, A. (1995). Analysis of the protein kinase activity of moss phytochrome expressed in fibroblast cell culture. Planta *197*, 592-596.

Valanne, N. (1966). The germination of moss spores and their control by light. Annals Botanici Fennici *3*, 1-40.

Zeidler, M., Lamparter, T., Hughes, J., Hartmann, E., Remberg, A., Braslavsky, S., Schaffner, K., & Gärtner, W. (1998). Recombinant phytochrome of the moss *Ceratodon purpureus*: heterologous expression and kinetic analysis of Pr-->Pfr conversion. Photochem.Photobiol. *68*, 857-863.

BLUE/UV-A LIGHT SIGNALING IN MOSS

EDWARD B. TUCKER

Department of Natural Sciences, Baruch College, City University of New York, 1 Bernard Baruch Way, New York, NY, 10010, USA

Abstract. Seed plants contain four identified photoreceptors and at least one unidentified photoreceptor for blue/UV-A light. Cryptochromes (CRY1 and CRY2) modulate growth while phototrophins (PHOT1 and PHOT2) modulate trophic responses to light. Flavins are the chromophores and absorb in the blue and UV-A range. Considerable information is known about the structure and mode of action of these photoreceptors. The moss *Physcomitrella patens* contains two identified cryptochromes (CRY1a and CRY1b) which modulate side branch formation and auxin metabolism. While moss responds in a phototrophic manner to blue/UV-A light, published reports identifying and characterizing phototropins have not yet appeared in the literature. Transduction of the blue/UV-A stimulus does involve calcium signaling in moss cells.

1. INTRODUCTION

Plants absorb and respond to blue (390-500nm)/UV-A (320-390nm) light through a network of signaling pathways that decode and transduce information within cells. Calcium plays a role in this signal transduction (Russell et al. 1998, Baum et al. 1999, Fig. 1 below). Down stream responses of moss to blue/UV-A light include: chloroplast movements (Kadota et al., 2000), polarity establishment (Cove and Lamparter, 1998), auxin regulation (Imaizumi et al. 2002), gravitropic response alterations (Demkiv et al. 1998, Khorkavtsiv and Kardash 2001, Young and Sack 1992) side branch induction (Imaizumi et al. 2002, Table 1 below), gametophore differentiation (Imaizumi et al. 2002) and gametophore growth regulation (Imaizumi et al. 2002). Previous reviews that include the topic of moss response to blue light have been published (Wada and Kadota 1989, Hartmann and Weber 1990, Christianson 2000).

In *Arabidopsis*, four blue light receptors have been identified in the cell: cryptochrome 1 (CRY1) (Ahmad and Cashmore 1993), cryptochrome 2 (CRY2)

A.J. Wood et al. (ed.),
New Frontiers in Bryology: Physiology, Molecular Biology and Functional Genomics, 177–188.

(Hoffman et al. 1996, Lin et al. 1996), phototropin 1 (PHOT1) (Huala et al. 1997, Christie et al. 1998, 1999) and phototropin 2 (PHOT2) (Jarillo et al. 1998). A UV-A photoreceptor, in addition to cryptochrome and phototropin, has been reported (Young et al. 1992) and a phytochrome-phototropin chimeric photoreceptor has been identified in the fern *Adiantum* (Nozue et al. 1998). In *Arabidopsis*, CRY1 and CRY2 can be phosphorylated by a phytochrome A-associated kinase (Ahmad et al. 1998a). The site of phosphorylation is within the C-terminal region of both CRY1 and CRY2. To attain phosphorylation, red light activates phytochrome which in turn activates a kinase, which phosphorylates CRY1 or CRY2. With exposure to blue light, CRY1 or CRY2 is fully activated. If only blue light is present, phytochrome will be only partially activated (Ahmad et al. 1998b).

Moss plants contain two identified cryptochrome photoreceptors (PpCRY1a and PpCRY1b) (Imaizumi et al. 2002) and are expected to contain phototropin(s) and possibly the unidentified UV-A photoreceptor and phytochrome-cryptochrome chimeric photoreceptor. *P. patens* cryptochromes are localized in the nucleus of light growing cells (Imaizumi et al. 2002) and perhaps elsewhere in the cytoplasm in dark growing cells. Phototropin are likely integrated within the plasma membrane as they are in *Arabidopsis* and a fraction may be released into the cytoplasm in response to blue light (Sakamoto and Briggs 2002). The chromophores of cryptochromes and phototropins are flavins, which absorb blue and UV-A light (Lin et al. 1995, Christie et al. 1998, Malhotra et al. 1995, Salomon et al. 2000, Kasahara et al. 2002, Ahmad et al. 2002). It appears that the absorption spectrum correlates with the action spectrum of some responses (Karlsson 1986, Quinones et al. 1996, Eisinger et al. 2000, Kinoshita et al. 2001) but not with others (Ahmad et al. 2002).

Moss cryptochromes belongs to a family that includes: (1) class I cyclobutane pyrimidine dimers (CPD) photolyase, (2) Class II CPD photolyase, (3) plant CRY and (4) (6-4)photolyase/animal CRY (Todo 1999, Deiserhofer 2000). Plant cryotpchromes have amino acid sequence that are similar to DNA photolyases except that cryptochromes have no photolyase activity and contain a C-terminal extension not found in photolyases. The C-terminal extension of cryptochromes differs and contain all of the sequence information necessary for signaling (Yang et al. 2000).

The phototropin family is plasma membrane-associated flavoproteins ranging from 114 to 130 kDa, that contains two (circa 110 amino acids) flavin-binding domains, LOV1 and LOV2, within the N terminus and a typical serine-threonine protein kinase domain at its C terminus (Briggs and Huala 1999, Briggs et al. 2001, Swartz et al. 2001). The LOV (light, oxygen, and voltage) domains belong to the PAS domain (characterized in PER, ARNT, and SIM proteins) superfamily of sensor proteins (Taylor and Zhulin 1999).

2. MOSS CRYPTOCHROME

P. patens contains two cryptochrome genes that have almost identical nucleotide sequences except for the 3' regions of the predicted 3' untranslated region; therefore, they were designated PpCRY1a and PpCRY1b (Imaizumi et al. 2002). The genomic DNA sequences corresponding to each cDNA revealed that these cDNAs are derived from two different genes. Southern blot analysis confirmed two CRY genes (Imaizumi et al. 2002). The CRY cDNAs from *P. patens* have an open reading frame of 2184 bp and encode proteins with 727 deduced amino acids and a deduced molecular mass of 82 kD (Imaizumi et al. 1999). A single base difference leads to PpCRY1a having threonine and PpCRY1b having methionine at amino acid 80. Three fern cryptochromes (Adiantum CRY1, CRY2 and CRY3) showed the highest homologies to the deduced amino acid sequence of this gene (71.7%, 73.5%, and 72.2%). The occurrence of the TGYP motif and residues known to interact with FAD (Park et al. 1998) in the N-terminus indicate that FAD is the likely chromophore. Cryptochrome of *Arabidopsis* bind FAD and MTHF, lack photolyase activity and regulate blue light elongation (Christie and Briggs 2001). There are similarities of PpCRY1 amino acid sequences in the C-terminus (DQLVP and STAESAS) to the conserved motifs (DQMVP and STAESSSS, respectively) found in the C-terminus of plant cryptochromes (Kangeae and Wada 1998). It has been reported that STAESSSS motif is important for the phosphorylation of cryptochrome by phytochrome A *in vitro* (Ahmad et al. 1998b) which can be taken to mean that moss CRY can be phosphorylated. These findings suggest that cryptochromes in both *Physcomitrella* and *Arabidopsis* share similar signal transduction cascades, in which phytochromes might be involved, via the C-terminal region.

The PpCRY1a and PpCRY1b are localized in the nucleus of protoplasts isolated from protonematal tissue growing under light. In this tissue, the gene products possess the putative nuclear localization signals of *Adiantum* cryptochrome (Imaizumi et al. 2000), and the (GUS)-PpCRY1a and (GUS)-PpCRY1b fusion proteins localized to the nucleus of protoplasts (Imaizumi et al. 2002). However, there is enough evidence of light-dependent cryptochrome degradation in *Arabidopsis* (Lin et al. 1998, Guo et al. 1999) and light dependent intracellular distribution of cryptochrome in *Adiantum* (Imaizumi et al. 2000) to speculate that PpCRY1 or PpCRY2 may be located in the cytoplasm as well as the nucleus in dark growing caulonemata.

3. MOSS PHOTOTROPIN

There are no published reports of isolated phototropin for moss. However, since the identified phototropins in higher plants mediate chloroplast relocation, unidentified phototropins in moss likely mediate the reported chloroplast relocations in moss (Kadota et al. 2000). And since phototropins are thought to be

responsible for the blue light induced calcium waves observed in *Arabidopsis* (Baum et al. 1999, Shimazaki et al. 1999), they probably are responsible for the calcium waves we have observed in *Physcomitrella* (Fig 1). While phototropin has not been identified in ferns, a chimeric phytochrome-phototropin protein (PHY3) has been reported (Kadota et al. 2000). The COOH –terminal domain of PHY3 possesses two 103-aa LOV domain repeats (44% identity) found in LOV1 of *Arabidopsis*, followed by a Ser/Thr PK domain (Nozue et al. 1998). The absorption spectra of the LOV (light, oxygen or voltage) domains of PHOT1 and PHOT2 (Sakai et al. 2001, Christie et al. 1999) that function as binding sites for the flavin mononucleitide chromophore closely match the action spectrum for stomatal opening. Thus, indirect evidence suggests phototropins in moss.

4. MODE OF ACTION

While the mechanism by which blue and UV-A light is converted by cryptochrome into a chemical signal is unknown, it may have similarities to the mode of action of its close relative photolyase (Todo 1999, Deisenhofer 2000). Photolyases contain reduced flavin adenine dinucleotide ($FADH_2$) and either methenyltetrahydrofolate (MTHF) or deazaflavin (DF) (Todo 1999, Johann and Deiserhofer 2000). Muller-Enoch (1997) demonstrated that a photoacceptor is able to transfer its excitation energy, λ_E =380 nm, and its fluorescence energy, λ_F= 460 nm, radiationless to the fluorophores FAD and FMN (λ_E=465nm; λ_F=525nm). In photolyase, MTHF or DF function as an antenna that becomes elevated into an energetically excited state after absorbing blue light. Energy is transferred from MTHF or DF to $FADH_2$, which donates an electron to the CPD substrate held in the enzyme's active site. This negative charge induces splitting of the cyclobutane ring of CPD. The electron is then transferred back to flavin, restoring the CPD into its original state. This electron transfer reaction can occur when the second chromophore is depleted from CPD photolyase. In this case, $FADH_2$ itself absorbs blue light. FAD is required for the DNA binding site and possesses a role in the protein structure. MacFarland and Stanley (2001) scanned photolyase and reported an absorption maximum at blue and UV-A regions. The action spectrum for blue-light- activated responses resembled the absorption spectrum of flavoproteins, with bands of activity in blue and UV-A region of the spectrum (Briggs and Olney 2001).

Phototropins undergo autophosphorylation in response to blue light. *E, coli*-expressed LOV domains bind riboflavin-5'-monophosphate, are photochemically active, and have major absorption peaks at 360 and 450 nm. Blue light excitation of the LOV2 domain generates a transient 660 nm-absorbing state that subsequently decays into a 390 nm-absorbing metastable form. The LOV2 domain (450 nm) recovers spontaneously. Phosphorylation occurs on serine and threonine residues, requires Mg^{2+} but not Ca^{2+}, and is highly ATP specific (Short and Briggs 1994).

5. CAULONEMA, THE BLUE/UV-A RECEPTOR CELL. DARK-GROWING CAULONEMATA ARE DIFFERENT THAN LIGHT-GROWING CAULONEMATA

Cultures of moss are started either as germinating spores, inoculum or homogenized tissue. These tissue fragments are placed either directly into liquid culture or are planted directly onto solid agar culture media. The first filaments to develop are chloronematal and later develop into caulonematal filaments. When mats of this tissue are placed into the dark, they develop dark growing caulonematal filaments. These caulonemal cells are morphologically and physiologically different from caulonemal cells produced in the light (Duckett et al. 1998). Side branches will develop off of the filaments of either light or dark growing caulonemata.

Dark growing caulonemal apical cells are polarized (Cove 2000) and contain in order: apical cap, plastid, nuclear, and basal vacuolar regions. These cells are tip-growing and like pollen, are presumed to contain a tip focused Ca^{2+} gradient that is maintained by an influx of Ca^{2+} through channels in the apex (Holdaway-Clarke et al. 1997, Hepler 1997, Yang 1998). They respond to changes in the direction of gravity (Cove and Knight 1987, Knight and Cove 1991, Knight et al. 1991) and exposure to light (Cove et al. 1978, Russell, et al. 1998). Exposure to blue light inhibits this gravi-response (Young and Sack 1992, Khorkavtsiv and Kardash 2001, Demkiv et al. 1998).

Chloroplasts in light growing caulonemata move in response to blue light (Kadota et al. 2000). When irradiated with polarized blue light (λ_{max} 452.5), chloroplast distribution in the cell depended upon the direction of the light. When light was parallel to cross-walls (perpendicular to filament) chloroplasts accumulated along the cross-wall; however no accumulation along the cross-wall was observed when the light was perpendicular to it (parallel to the filament axis). When a part of the cell was irradiated with a microbeam of blue light, chloroplasts accumulated at or avoided the illumination point depending on the fluence rate used. Low fluence rate (0.01-85.5 W m^{-2}) induced accumulation while high fluence rate (285 W m^{-2}) induced avoidance. Chloroplasts in *P. patens* could move either along microtubules or microfilaments in response to either 10 W m^{-2} or 100 W m^{-2} blue light (λ_{max} 452.5) but only along microtubules, in response to (low or high fluence) red light (Sato et al. 2001a). While external calcium was required for movement induced by touch, it was not required for blue light (λ_{max} 452.5 with half-band width 7.5) stimulated movement in *Adiantum* (Sato et al. 2001b).

The morphological events of side branch formation are complex (Schumaker and Dietrich 1997, Schmiedel and Schnept 1979) and indicate a series of physiological changes that we know nothing about. The steps include a small swelling at the apical end of the cell, elongation of the swelling, nuclear migration

to this area, division and then partition of one daughter nucleus to the branch initial and migration of the other one back midway along the parent cell. Finally a cell wall forms parallel to the longitudinal axis of the filament cell, to separate the initial cell from the filament. In the last step, this side-branch initial continues to develop in the dark into a side-branch. Buds, which develop into gametophores will develop off of the side branches. Both blue and UV-A light stimulate side branch formation but not bud formation. Bud formation can be induced by red light (Mitra et al. 1965) or treatment with cytokinins (Schumaker and Dietrich 1998). Exposure of dark-adapted caulonemata to red light induced a Ca^{2+} dependent transient membrane depolarization and side branch formation (Johannes et al. 1997).

The growth of protonemal colonies of wild type and cryptochrome disruptants was observed under various light conditions (Imaizmi et al. 2002). The appearance of Ppcry1a, Ppcry1b and Ppcry1acry1b colonies were indistinguishable from those of wild-type colonies under 10 days of continuous white or red light (26 μmoles $m^{-2}s^{-1}$), whereas a clear difference was observed under blue light (28 μmoles $m^{-2}s^{-1}$). With blue light, the colony diameter of each disruptant was larger than that of the wild type, and the densities of the protonemata in each colony decreased in order from wild type, to Ppcry1a, to Ppcry1b to Ppcry1acry1b. The difference was a result of lack of side branch formation on the protonemata of the Ppcry knock-outs. Ppcry1acry1b formed only very few side branches under constant blue light. In addition, while the red light (26 μmoles $m^{-2}s^{-1}$) effect was significantly reduced by the simultaneous irradiation with far-red light (12 μmoles $m^{-2}s^{-1}$), blue light (28 μmoles $m^{-2}s^{-1}$) induction of branches were not affected at all with far-red light, indicating that blue light receptor other that phytochrome is responsible for the blue light response. (Kadota et al. 2001).

	Wild Type		Ppcry1acry1b	
	Initials Ave(SD)	Side-Branches Ave(SD)	Initials Ave(SD)	SideBranches Ave(SD)
Dark	0.36(0.05)	0.20(0.03)	0.40(0.05)	0.15(0.04)
UVA	1.52(0.11)	0.73(0.07)	1.40(0.11)	0.66(0.06)
Blue	0.78(0.08)	0.47(0.07)	0.79(0.10)	0.42(0.03)

Table 1 *Initials and side-branches were formed on dark-growing caulonemata of Wt and Ppcry1acry1b. Moss was grown 8 days in light, 8 days in dark to produce caulonemata, exposed 2 min blue (20 μW cm^{-2}) or UV-A light (20 μW cm^{-2}), returned to darkness for 2 days, and then observed on a dissection microscope Sample number is 150.*

The formation of side branches and side branch initials on dark growing caulonemata of either wild type or Ppcry1acry1b induced by either blue or UV-A light was compared in our lab. A 2 min exposure of 20 μW cm^{-2} of either blue (λ_{max} 488 nm) light or UV-A (λ_{max} 365 nm) light onto dark-growing caulonemata was done. Both the wild type and double-knockout had very few initials or side branches when growing in the dark (i.e., controls). In the wild type, the blue light induced an average of 0.8 initials and 0.5 side branches per filament while the UV-A induced 1.5 initials and 0.7 side branches (Table 1). In Ppcry1acry1b the blue light induced 0.8 initials and 0.4 side branches while UV-A induced 1.4 initials and 0.7 side branches. Thus, in dark growing caulonemata, UV-A was more effective in the induction of initials and side branches than was blue light and there was little difference between the wild type and Ppcry1acry1b. In conclusion, while cryptochrome plays a major role in side branch formation in light-growing protonemata (Imaizumi et al. 2002), our data illustrate that it does not play a similar role in side branch formation on dark-growing caulonemata. Because UV-A light was more effective than blue light at side branch formation, we speculate that UV-A has an additional physiological pathway. UV-A has been reported to regulate chalcone synthase gene expression in *Arabidopsis* (Fuglevand et al. 1990, Long and Jenkins 1998; Jenkins et al. 2001).

6. SIGNALING/CALCIUM

When Russell et al. (1998) illuminated protonemal tissue of dark grown transgenic *P. patens* expressing the apoaequorin gene, with blue light (λ_{max} 450 nm) for 3 s at 100 μmol m^{-2} s^{-1}, the light induced a Ca^{2+} transient of between 200 to 700 nM that lasted 30 to 40 s. There was a delay of 6 s between illumination and the beginning of the $[Ca^{2+}]_{cyt}$ increase, and Ca^{2+} in the medium was not required. A 1 min exposure to blue light resulted in the production of side branches from subapical cells of dark grown caulonemata (Russell et al. 1998). Results from experiments using single *cry1*, *cry2* and *phot1* single knockouts in transgenic apoaequorin *Arabidopsis*, indicate that the blue light induced $[Ca^{2+}]_{cyt}$ transients is a response from phototropin and not cryptochrome (Baum et al., 1999).

Ca^{2+} - ratio imaging using Fura-2 (λ_{max} 340 nm-Ca^{2+} dependent & λ_{max} 380 nm- Ca^{2+} independent) was undertaken in our lab to study $[Ca^{2+}]_{cyt}$ in UV- A light signaling in Wt and Ppcry1acry1b. Fortuitously, the UV-A light used to measure $[Ca^{2+}]_{cyt}$ also stimulated the calcium response. A single caulonemal cell that was growing in the dark was used as compared to the mat of protonema used by Russell et al. (1998). And UV-A was used instead of blue light. Cells in growth media were exposed repeatedly with 1 s exposure with intensity of 30 μW cm^{-2}. Fura-2 ratio images of the apical cell were obtained immediately following UV-A exposure and demonstrated two intracellular waves (Fig 1). One $[Ca^{2+}]_{cyt}$ was a "tip wave," i.e. a wave which originates from a discrete area at the apex of the cell

in the apical cap and moves towards the base of the cell. The $[Ca^{2+}]_{cyt}$ peaked circa 600 – 800 nM and the waves oscillated multiple times. A second type $[Ca^{2+}]_{cyt}$ response was a "basal wave," i.e. a wave which originates in the area adjacent to the cell junction and moves towards the tip of the cell. The $[Ca^{2+}]_{cyt}$ peaked circa 800 – 1200 nM and the waves oscillated multiple times. These Ca^{2+} waves traveled at an average rate of 16 $\mu m.s^{-1}$, and each wave was completed circa 20 s. The data indicate that dark grown caulonemal apical cells of both Wt and Ppcry1acry1b contain photoreceptors, which upon stimulation with UV-A light induce two different $[Ca^{2+}]_{cyt}$ transients.

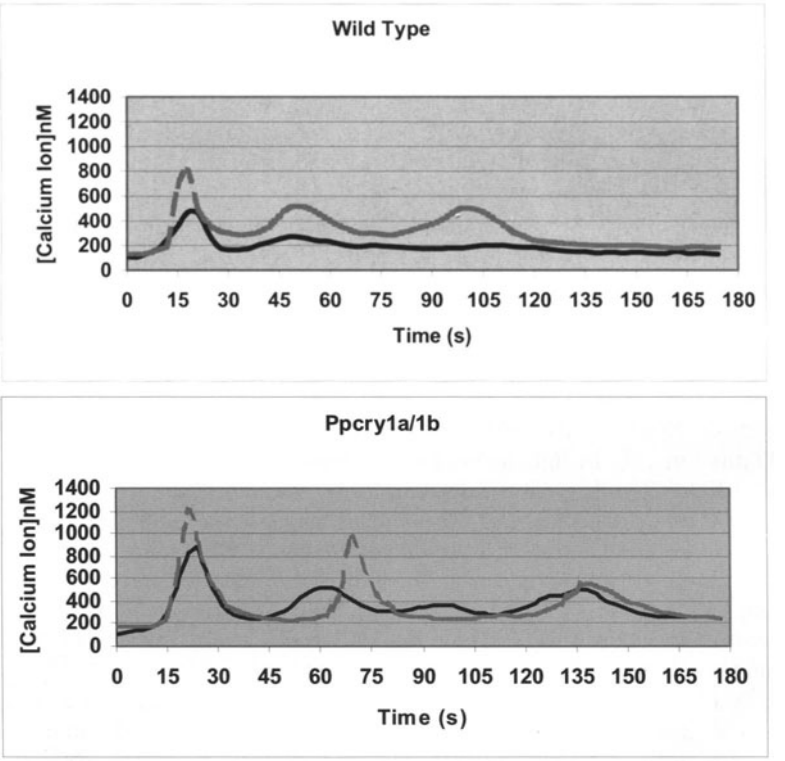

Figure 1. λ_{max} 380 & 340 induced $[Ca^{2+}]_{cyt}$ tip and base region waves in P. patens Wt and Ppcry1acry1b caulonemata. $[Ca^{2+}]_{cyt}$ at growing tip and basal region calculated from ratio images. Images taken every 3 seconds for 174 seconds. The solid line is the tip wave traveling in a proximal direction from apical cap. The broken gray line is the basal wave traveling in a distal direction from the cytoplasm near the cell junction. Multiple oscillations are observed.

7. CONCLUSIONS AND PROSPECTS

An understanding of how plants perceive, transduce and then respond to blue/UV-A light is still in its early stages. *P. patens* gene mutagenesis by gene targeting enables the direct generation of loss-of-function and point mutations in *P. patens CRY*, *PHOT* and other genes involved in blue/UV-A light signaling pathways. These tools will help delineate the steps and mechanisms of blue/UV-A light signaling. Gene targeting is still unfeasible as a routine procedure in plants except for *P. patens* (Schaefer and Zryd 2001). Preliminary analysis of ESTs and of genomic sequences clearly indicates that the genes of *P. patens* and other land plant are highly similar, at the level of both intron-exon structure and codon usage. Several labs in America, Japan and Europe are now using moss to understand light signaling.

Acknowledgement

NASA (NAG5-3743) and NSF (IBN-0214687) have supported my work on blue/UV-A light signaling in moss. I thank Drs. Imaizumi, Kadota, Haseby and Wada for use of *P. patens* cryptochrome knock-outs and Shaan Alli and Steven Davidowitz for research assistance.

8. REFERENCES

Ahmad, M., & Cashmore, A.R. (1993) *HY4* gene of *A. thailana* encodes a protein with the characteristics of a blue-light photoreceptor. *Nature*, 366, 162-166.

Ahmad, M., Grancher, N., Heil, M., Black, R.C., Giovani, B., Galland, P., & Lardemer, D. (2002) Action spectrum for cryptochrome-dependent hypocotyl growth inhibition in *Arabidopsis*. *Plant Physiol.*, 129, 774-785.

Ahmad, M., Jarillo, J.A., & Cashmore, A.R. (1998a). Chimeric proteins between cry1 and cry2 *Arabidopsis* blue light phororeceptors indicate overlapping functions and varying protein stability. *Plant Cell,* 10, 197-207.

Ahmad, M., Jarillo, J.A., Smirnova, O., & Cashmore, A.R. (1998b). The CRY1 blue light phororeceptor of Arabidopsis interacts with phytochrome A in vitro. *Mol. Cell*, 1, 939-948.

Baum, G., Long, J.C., Jenkins, G.I., & Trewavas, A.J. (1999) Stimulation of the blue light phototropic receptor NPH1 causes a transient increase in cytosolic Ca^{2+}. *Proc. Natl. Acad. Sci.*, 96, 13554-13559.

Briggs, W.R., Beck, C.F., Cashmore, A.R., Christie, J.M., Hughs, J., Jarillo, J.A., Kagawa, T., Kanegae, H., Liscum, E., Nagatani, A., Okada, K., Salomon, M., Rudiger, W., Sakai, T., Takano, M., Wada, M., & Watson, J.C. (2001) The phototropin family of photoreceptors. *The Plant Cell*, 13, 993-997.

Briggs, W.R. & Huala, E. (1999) Blue-light photoreceptors in higher plants. *Annu. Rev. Cell Dev. Biol.*, 15, 33-62.

Briggs, W.R., & Olney, M.A. (2001) Photoreceptors in plant morphogenesis to date. Five phytochromes, two cryptochromes, one phototropin, and one superchrome. *Plant Physiol.*, 125, 85-88.

Christianson, M.L. (2000) Control of morphogenesis in bryophytes. In "Bryophyte Biology" (Shaw, A.J. & Goffinet, B, eds.) Cambridge Univ. Press Cambridge, UK

Christie, J.M., & Briggs, W.R. (2001) Blue light sensing in higher plants. *J. Biol. Chem.*, 276, 11457-11460.

Christie, J.M., Reymond, P., Powell, G.K., Bernasconi, P., Raibekas, A.A., Liscum, E., & Briggs, W. (1998) *Arabidopsis* NPH1: a flovoprotein with the properties of a photoreceptor for phototropism. *Science*, 282, 1698-1701.

Christie, J.M., Salomon, M., Nozue, K., Wada, M., & Briggs, W.R. (1999) LOV (light, oxygen, or voltage) domains of the bluc light photoreceptor phototropin (nph1): Binding sites for the chromophore flavin mononucleotide. *Proc. Natl. Acad. Sci. USA* 96, 8779-8783.

Cove DJ (2000) The generation and modification of cell polarity. *J. Exp. Bot.* 51, 831-838

Cove, D.J., Lamparter, T. (1998) The role of light in the regulation of moss development. Microbial Responses to light and time. In "Soc. Gen. Microb. Symp. 56" (Caddick, M.X., Baumgerg, S., Hodgson, D.A., & Phillips-Jones, M.K., eds) Cambridge Univ. Press

Cove, D.J., & Knight, C.D. (1987) Gravitropism and phototropism in the moss *Physcomitrella patens*. In "Developmental Mutants of Higher Plants" (Thomas, H. & Grierson, D. eds). Cambridge Univ. Press, Cambridge, UK.

Cove, D.J., Schild, A., Ashton, N.W., & Hartmann, E. (1978) Genetic and physiological studies of the effect of light on development of the moss *Physcomitrella patens*. *Phytochem. and Phytobiol.*, 27, 249-254

Deisenhofer, J. (2000) DNA photolyases and cryptochromes. *Mutation Research* 460, 143-149.

Demkiv, O.Y., Kordyum, E.L., Khorkavtsiv, Y.D., Kardash, O.R., & Chaban, C.I. (1998) Gravi-and photostimuli in moss protonema growth movements. *Adv. Space Res.*, 21, 1191-1195.

Duckett, J.G., Schmid, A.M., & Ligrone, R. (1998) Protonemal morphogenesis. In "Bryology for the Twenty-first Century" (Bates, J.W., Ashton, N.W., & Duckett, J.G. eds.) British Bryological Society, UK

Eisinger, W., Swartz, T.E., Bogomolni, R., & Taiz, L. (2000) The ultraviolet action spectrum for stomatal opening in broad beans. *Plant Physiol.*, 122, 99-105.

Fuglevand, G., Jackson, J.Q., Jenkins, & G.I. (1996) UV-B, UV-A, and blue light signal transduction pathways interact synergistically to regulate chalcone synthase gene expression in *Arabidopsis*. *The Plant Cell*, 8, 2347-2357.

Guo, H., Duong, H., Ma, N., & Lin, C. (1999) The *Arabidopsis* blue light receptor cytochrome 2 is a nuclear protein regulated by a blue light-dependent post-transcriptional mechanism. *Plant J.* 19, 279-287.

Hartman, E., & Weber, M. (1990) Photomodulation of protonema development. In "Bryophyte development: physiology and biochemistry" (Chopra, R.N., & Bhatla, S.C. eds). CRC Press, Inc Boca Raton, FL, USA

Hoffman, P.D., Batschauer, A., & Hays, J.B. (1996) *PHH1*, a novel gene from *Arabidopsis thaliana*, that encodes a protein similar to plant blue-light photoreceptors and microbial photolyases. *Mol. Gen. Genet.*, 253, 259-265.

Hepler, P.K. (1997) Tip growth in pollen tubes: calcium leads the way. *Trends in Plant Sci.*, 2, 79-80

Holdaway-Clarke, T.L., Feijo, J.A., Hackett, G.R., Kunkel, J.G., & Hepler, P.K. (1997) Pollen tube growth and the intracellular cytosolic calcium gradient oscillate in phase while extracellular calcium influx is delayed. *The Plant Cell*, 9, 1999-2010

Huala, E., Oeller, P.W., Liscum, E., Han, I-S., Larsen, E., & Briggs, W.R. (1997) *Arabidopsis* NPH1: A protein kinase with a putative redox-sensing domain. *Science*, 278, 2120-2123.

Imiazumi, T., Kadota, A., Haseby, M., & Wada, M. (2002) Cryptochrome light signals control the development to suppress auxin sensitivities in the moss *Physcomitrella patens*. *The Plant Cell*, 14, 373-386.

Imaizumi, T., Kanegae, T., & Wada, M. (2000) Cryptochrome nucleocytoplasmic distribution and gene expression are regulated by light quality in fern *Adiantum capillus-veneris*. *The Plant Cell*, 12, 81-95.

Imaizumi, T., Kiyosue, T., Kanegae, T., & Wada, M. (1999) Cloning of the cDNA encoding blue-light photoreceptor (cryptochrome) from the moss *Physcomitrella patens* (Accession No. AB027528) *Plant Physiol.*, 120, 1205

Jarillo, J.A., Ahmad, M., Cashmore, A.R. (1998) NPL1 (accession No. AF053941): A second member of NPH serine/threonine kinase family of *Arabidopsis*. *Plant Physiol.*, 117, 719.

Jenkins, G.I., Long, J.C., Wade, H.K., Shenton, M.R., & Bibikova, T.N. (2001) UV and blue light signaling: pathways regulating chalcone synthase gene expression in *Arabidopsis*. *New Phytologist*, 151, 121-131.

Johannes, E., Ermolayeva, E., & Sanders, D. (1997) Red light-induced membrane potential transients in the moss *Physcomitrella patens*: ion channel interaction in phytochrome signaling. *J. Exp. Bot.*, 48, 599-608.

Kadota, A., Imaizumi, T., Hasebe, M., & Wada, M. (2001) Cryptochrome-dependent branch induction in Physcomitrella patens. In "Moss 2001; An international meeting on moss biology" (Shin, J. S, & Hasebe, M. eds) Okazaki, Japan.

Kadota, A., Sato, Y., & Wada, M. (2000). Intracellular chloroplast photorelocation in the moss *Physcomitrella patens* is mediated by phytochrome as well as by a blue-light receptor. *Planta*, 210, 932-937.

Kanegae, T., & Wada, M. (1998) Isolation and characterization of the plant blue light photoreceptor (cryptochrome) homologous genes of the fern *Adiantum capillus-veneris*. *Mol. Gen. Genet.*, 259, 345-353.

Karlsson, P.E. (1986) Blue light regulation of stomata in wheat seedlings. II Action spectrum and search for action dichroism. *Physiol. Plant.*, 66, 207-210.

Kasahara,M., Swartz, T.E., Olney, M.A., Onodera, A., Mochizuki, N., Fukuzawa, E., Tabata, S., Kanegae, H., Takano, M., Christie, J.M., Nagatani, A., & Briggs, W.R. (2002) Photochemical properties of the flavin mononucleotide-binding domains of the phototropins from *Arabidopsis*, rice and *Chlamydomonas reinhardtii*. *Plant Physiol.*, 129, 762-773.

Khorkavtsiv, O.Y., & Kardash, O.R. (2001) Gravi-dependent reactions of the moss *Pohlia nutans* protonemata. *Adv. Space Res.*, 27, 989-993.

Kinoshita, T., Doi, M., Suetsugu, N., Kagawa, T., Wada, M., & Shimazaki, K. (2001) Phot1 and phot2 mediate blue light regulation of stomatal opening. *Nature*, 414, 656-660.

Knight, C.D., & Cove, D.J. (1991) The polarity of gravitropism in the moss *Physcomitrella patens* is reversed during mitosis and after growth on a clinostat. *Plant Cell and Env.*, 14, 995-1001

Knight, C.D., Futers, T.S., & Cove, D.J. (1991) Genetic analysis of a mutant class of *Physcomitrella patens* in which the polarity of gravitropism is reversed. *Mol. Gen. Genet.*, 230, 12-16

Lin, C., Ahmad, M., & Cashmore, A.R. (1996) *Arabidopsis* cryptochrome 1 is a soluble protein mediating blue light-dependent regulation of plant growth and development. *Plant J.*, 10, 893-902.

Lin, C., Robertson, D.E., Ahmad, M., Raibekas, A.A., Jorns, M.S., Dutton, P.L., & Cashmore, A.R. (1995) Association of Favin Adenine Dinucleotide with the *Arabidopsis* blue light receptor CRY1. *Science*, 269, 968-970.

Lin, C., Yang, H., Guo, H., Mockler, T., Chen, J., & Cashmore, A.R. (1998) Enhancement of blue-light sensitivity of *Arabidopsis* seedlings by blue light receptor cryptochrome 2. *Proc. Natl. Acad. Sci. USA*, 95, 2686-2690.

Long, J.C., & Jenkins, G.I. (1998) Involvement of plasma membrane redox activity and calcium homeostasis in the UV-B and UV-A/blue light induction of gene expression in *Arabidopsis*. *The Plant Cell*, 10, 2077-2086.

MacFarlane IV, A.W., & Stanley, R.J. (2001) Evidence of powerful substrate electric fields in DNA photolyase: implications for thymine dimmer repair. *Biochem.*, 40, 15203-15214.

Malhotra, K., Sang-Tae, K., Batschauer, A., Dawut, L., & Sancar, A. (1995) Putative blue-light photoreceptors from *Arabidopsis thaliana* and *Sinapis alba* with a high degree of sequence homology to DNA photolyase contain the two photopyase cofactors but lack DNA repair activity. *Biochem.*, 34, 6892-6899.

Mitra, G.C., Misra, L.P., & Prabha, C. (1965) Interaction of red and blue light on the development of the protonema and bud formation in *Pohlia nutans*. *Planta*, 65, 42

Muller-Enoch, M. (1997) Blue light mediated photoreduction of the flavoprotein NAHPH-cytochrome P450 reductase. A Forster-type energy transfer. *Zeitschrift fur Naturforschung J. Biosci.*, 52, 605-614.

Nozue, k., Kanegae, T., Imaizumi, T., Fukuda, S., Okamoto, H., Yeh, K-C., Lagarias, J.C., & Wada, M. (1998). A phytochrome from the fern *Adiantum* with features of the putative photoreceptor NPH1. *Proc. Natl. Acad. Sci. USA*, 95, 15826-15830.

Park, H-W., Kim, S-T., Sancar, A., & Deisenhofer, J. (1998) Crystal structure of DNA photolyase from *Escherichia coli*. *Science 268*, 1866-1872.

Quinones, M.A., Lu, Z., & Zeiger, E. (1996) Close correspondence between the action spectra for the blue light responses of the guard cell and coleoptile chloroplasts, and the spectra for blue l;ight-dependent stomatal opening and coleoptile phototropism. *Proc. Natl. Acad. Sci.*, 93, 2224-2228.

Russell, A.J., Cove, D.J., Trewavas, A.J., & Want, T.L. (1998) Blue light but not red light induces a calcium transient in the moss *Physcomitrella patens* (Hedw.) B., S. &G. *Planta*, 206, 278-283

Salomon, M., Christie, J.M., Knieb, E., Lempert, U., & Briggs, W.R. (2000) Photochrmical and mutational analysis of the FMN-binding domain of the blue-light photoreceptor phototropin. *Biochem.*, 39, 9401-9410.

Sakai, T., Kagawa, T., Kasahara, M., Swartz, T.E., Christie, J.M., Briggs, W.E., Wada, M., & Akada, K. (2001) *Arabidopsis* nph1 and nph11: Blue light receptors that mediate both phototropism and chloroplast relocation. *Proc. Natl. Acad. Sci. USA*, 98, 6969-6974.

Sakamoto, K., & Briggs, W.E. (2002) Cellular and subcellular localization of phototropin-1. *Plant Cell*, 14, 1723-1735.

Sato, Y., Wada, M., & Kadota, A. (2001)a Choice of tracks, microtubules and/or actin filaments for chloroplast photo-movement is differentially controlled by phytochrome and a blue light receptor. *J. Cell Sci.*, 114, 269-279.

Sato, Y., Wada, M., & Kadota, A. (2001)b External Ca^{2+} is essential for chloroplast movement induced by mechanical stimulation but not by light stimulation. *Plant Physiol.*, 127, 497-504.

Schaefer, D., & Zryd, J-P. (2001) The moss *Physcomitrella patens*, now and then. *Plant Physiol,*. 127, 1430-1438.

Schmiedel, G., & Schnipf, E. (1979) Side branch formation and orientation in the caulonema of moss, *Funaria hygrometrica*: Normal Development and fine structure. *Protoplasma*, 100, 367-389.

Schumaker, K.S., & Dietrich, M.A. (1997) Programmed changes in form during moss development. *The Plant Cell*, 9, 1099-1107.

Schumaker, K.S., & Dietrich, M.A. (1998) Hormone-induced signaling during moss development. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 49, 501-523.

Shimazaki, K., Goh, C.H., & Kinoshita, T. (1999) Involvement of intracellular Ca^{2+} in blue light-dependent proton pumping in guard cell protoplsts from *Vicia faba*. *Physiol. Plant.*, 105, 554-561.

Short, T.W., & Briggs, W.R. (1994) The transduction of blue light signals in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 45, 143-171.

Swartz, T.E., Corchnoy, S.B., Christie, J.M., Lewis, J.W., Szundi, I., Briggs, W.R., & Bogomolni, R.A. (2001) The photocycle of a flavin-binding domain of the blue light phororecepor phototropin. *J. Biol. Chem.*, 276, 36493-36500.

Taylor, B.L., & Zhulin, I.B. (1999) PAS domains: internal sensors of oxygen, redox potential, and light. *Microb. Mol. Biol. Rev.*, 63, 479-506.

Todo, T. (1999) Functional diversity of the DNA photolyase/blue light receptor family. *Mutation Research*, 434, 89-97

Wada, M., & Kadota, A. (1989) Photomorphogenesis in lower green plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 40, 169-191.

Yang, S. (1998) Signaling tip growth in plants. *Cur. Opin. Plant Biol.*, 1 525-530

Yang, H-Q., WU, Y-J., Tang, R-H., Liu, D., Liu, Y., & Cashmore, A.R. (2000) The C termini of *Arabidopsis* cryptochromes mediate a constitutive light response. *Cell*, 103, 815-827.

Young, J.C., Liscum, E., & Hangarter, R.P. (1992) Spectral-dependence of light-inhibiting hypocotyls elongation in photomorphogenic mutants of *Arabidopsis*: evidence for a UV-A photosensor. *Planta*, 188, 106-114.

Young, J.C., & Sack, F.D. (1992) Time-lapse analysis of gravitropism in *Ceratodon* protonemata. *Amer. J. Bot.*, 79, 1348-1358.

THE USE OF MOSSES FOR THE STUDY OF CELL POLARITY

DAVID J. COVE AND RALPH S. QUATRANO

Centre for Plant Sciences, University of Leeds, Leeds LS2 9JT, UK (DJC); Department of Biology, Washington University, One Brookings Drive, St. Louis, MO 63130, USA (RSQ)

Abstract. Moss development allows both the study of the development of cell polarity and the modification of polarity in an already-polar cell. The gametophore allows polarity studies to be extended to multicellular structures. Both spore germination and protoplast regeneration allow the study of the generation of a polar cell axis. The polarity of the axis of regenerating protoplasts of *Ceratodon purpureus* is influenced by light direction. The programming of the response is however complex. There is a delay before a response to a changed light direction is observed, indicating that axis polarity is fixed before asymmetrical development can be observed. However, the length of the delay is influenced by the state of the cell at the time the light direction is changed. When protoplasts regenerating in red light at 25°, are reoriented with respect to the light direction, there is a lag of about 9 hours before a response is observed. If protoplast are irradiated with far-red light immediately before reorientation, the lag is shorter, indicating that protoplasts use phytochrome to "memorize" light direction, preventing a precipitous response to temporarily changed conditions. Protonemal apical cells show tropic responses to both light and gravity. Mutant studies show that the phototropic response is mediated by phytochrome, and that this photoreceptor also turns off the gravitropic response in light. Mutants with a reversal of the orientation of their gravitropic response, have been isolated in both *Physcomitrella patens* and *C. purpureus*, and it is possible that in the latter species, a single gene can mutate either to prevent the gravitropic response or to reverse its orientation.

1. INTRODUCTION

We have defined polarity, in the context of developmental biology, as "the persistent, asymmetrical and ordered distribution of structures along an axis" (Cove et al., 1999). Moss development provides unrivalled opportunities to study polarity, not only at the level of the individual cell but also in multicellular organs. Both germinating spores and regenerating protoplasts provide material for the study of the acquisition of a polar axis by an apparently-unpolarised cell. The responses of protonemal apical cells to light and gravity, provide opportunities for the elucidation of the mechanisms by which the polarity of a cell's pre-formed axis can be modified in response to changed environmental inputs. The phototropic and gravitropic responses of gametophore allow similar studies to be made on a multicellular organ. The formation of three-dimensional buds from two-dimensional filaments allows the study

A.J. Wood et al. (ed.),
New Frontiers in Bryology: Physiology, Molecular Biology and Functional Genomics, 189–203.

of asymmetrical cell division during the transition from a two to a three dimensional structure. Furthermore, the haploidy of the gametophyte generation allows the direct isolation of mutants altered in aspects of cell and organ polarity, and simplifies their subsequent genetic analysis.

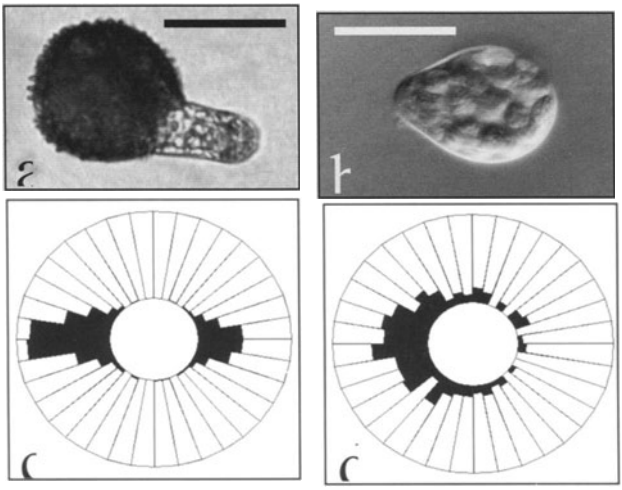

Figure 1 a. Germinating P. patens spore, scale bar 50 µm; b. C. purpureus regenerating protoplast, scale bar 50 µm; c, d. Distribution of axes of protoplasts of C. purpureus *regenerating in monochromatic c. red light (665 nm) and d. blue light (437 nm) (Cove et al, 1996)*

2. SPORE GERMINATION

Moss spores germinate by polar outgrowth, to form protonemal filaments that extend by tip growth of the apical cell and its serial (figure 1 a). The induction of spore germination in *Physcomitrella patens* requires light (Cove et al., 1978) but this does not hold for all moss species. Spores of *Funaria hygrometrica* germinate in darkness (Bauer and Mohr, 1959) but the point of emergence of the protonemal filament is sensitive to light (Jaffe and Etzold, 1965). In low intensities of red light, filament apices are produced towards the direction of the light source, whereas at high light levels, filaments emerge from the shaded side of the spore. The position that protonema emerge from the spore coat in *P. patens* is not however cued by the direction of light (Schild, 1981) so it appears that in this species, light does not play a role in determining the orientation of the polar axis formed during spore germination. Gravity also appears to play no role (Schild and Cove, unpublished data).

3. PROTOPLAST REGENERATION

Isolated protoplasts from a number of moss species have been shown to regenerate by polar outgrowth, resembling germinating spores (see figure 1 b). In

Ceratodon purpureus there is clear evidence that light plays a role in setting up the polar axis (Cove et al, 1996). By observing the polarity of the regeneration axes of populations of protoplast cultured in unidirectional monochromatic light, it has been established that in red light, axes are well aligned with the light direction, but only 65-75% of protoplasts show polar outgrowth towards the light source, the remainder growing away from it. In blue light, alignment with the direction of the light source is poorer but almost all axes are oriented towards the light source. Figure 1 c and d show circular histograms of the polarity of regeneration axes for these two light conditions.

Protoplast regeneration is not synchronous. At 25°, some protoplasts become asymmetrical as early as 15h after isolation, but the slowest protoplasts do not regenerate until 50 h after isolation. The time at which regeneration occurs in protoplasts left undisturbed in a unilateral light source, does not affect the alignment nor the orientation of the axis. Samples of protoplasts regenerating early or late show the same distribution of axis polarity. When protoplasts are re-oriented with respect to light direction during the course of regeneration, there is a lag of about nine hours (at 25° in monochromatic red light) before regeneration axes become aligned to the new light direction. As a result, it has been proposed (Cove et al, 1996) that the regeneration axis is fixed sometime before protoplasts become visibly asymmetrical. Reorientation affects not only axis alignment, but also whether outgrowth occurs towards or away from the light source (axis orientation). Soon after the light direction is changed, the proportion of protoplast orienting towards the light source declines. Although axes become aligned to the new light direction after a lag, differential orientation towards the light source does not occur, and equal numbers of protoplasts orient away and towards the light source. As a result of this, it has been proposed (Cove et al, 1996) that axis alignment and axis orientation may require different signal-transduction pathways.

Recent work (D.J.Cove, E.Hartmann, T.Lamparter, and R.S.Quatrano, unpublished) has investigated the role of phytochrome in determining the polarity of the protoplast regeneration axis. Samples of protoplasts which had started to regenerate in red light, were treated with far-red light, before being re-oriented and exposed to red light from a new direction. The far-red treatment results in the "memory" of the first light direction being lost more quickly, and a more rapid response to the new light direction, compared to the control (no far-red) treatment. However, if far-red light treated protoplasts are briefly returned to red light from the original direction before reorientation, the protoplasts' normal regeneration program is restored, *i.e.* protoplasts treated in this way resemble control protoplasts. These experiments provide evidence that phytochrome is involved in the perception of light direction. We further propose that the morphogen gradient responsible for the establishment of the polarity of the protoplast regeneration axis, must be stabilised by phytochrome in its P_{FR} form, resulting in the response to a new light direction being slow, but that the gradient is less stable when phytochrome is in its P_R form, allowing a response to a new light direction to occur more rapidly.

4. CHLORONEMAL APICAL CELL PHOTOTROPISM

In *P.patens*, primary chloronemata, *i.e.* the protonemal filaments that emerge from germinating spores, show a polar response to both unidirectional and polarised light (Jenkins & Cove, 1983a). The response is sensitive to both the wavelength and intensity of the light. In high intensities of red light, the polarotropic response of apical cells is to align their growth axis parallel to the electrical vector of the plane polarised light. The corresponding phototropic response is alignment perpendicular to the direction of the light source ("high intensity" response).

In low intensities of polarised red light, growth is aligned perpendicular to the E vector, while in low intensities of unidirectional light, the phototropic response is alignment parallel to the light direction, with most filaments growing towards the light source, but a few growing away ("low level" response). In red light, the change from the low to the high level response occurs in the range of 1 – 3 μmol quanta m^{-2} s^{-1}. At all wavelengths tested (range 417 to 730), both a polarotropic and a phototropic response occurred. In wavelengths between 540 and 600 nm and above 700 nm only a low level response was observed even at the highest fluence rates available. In wavelengths below 440 nm, only the high intensity response was observed even at t he lowest fluence rates at which growth occurred. However, higher irradiances of blue light tended to elicit growth away from the light source rather than the observed response to high irradiances of red light, of growth perpendicular to the light source. In other wavelengths, a transition from the low level to the high level response was observed, but not necessarily at the same fluence rate. Table 1 summarises these results.

These same studies (Jenkins & Cove, 1983a) also examined the possible role of phytochrome in the phototropic and polarotropic responses. When far red light (730 nm) was shone from above onto a Petri dish irradiated through the edge with a high intensity of red light (665 nm), the response to red light was similar to that observed for a much lower fluence rate of red light (*i.e.* in the absence of the far red irradiation). Thus phytochrome appears to be involved at least in the detection of the light intensity at which the switch from a low level to a high level response occurs.

A parallel study (Jenkins & Cove, 1983b), examined the phototropic and polarotropic responses of mutants that had been isolated as a result of their being deficient in the phototropic response of their caulonemal apical cells and their gametophores (see below). It was found that the primary chloronemal apical cells of these mutants still showed both phototropic and polarotropic responses. However, the responses differed from those of the wild type. Surprisingly, the mutants switched from the low level to the high level response at lower fluence rates. For example, in red light (665 nm), mutants showed a high level polarotropic response even in the lowest fluence rates tested (300 nmol quanta m^{-2} s^{-1}) . n green light (542 nm), the wild type showed a low level response at all intensities tested (maximum 70 μmol quanta m^{-2} s^{-1}), while the mutants switched for a low level to a high level phototropic response at fluence rates around 3 μmol quanta m^{-2} s^{-1}.

wavelength (nm)	intensity (μmol quanta m^{-2} s^{-1})
417	<1
442	0.5
473	1
498	3
542	>50
578	>50
613	30
637	4
665	1
687	1
715	>50
730	>50

Table 1. Relationship between fluence rate and polarotropic response of primary chloronemal apical cells of *P. patens*. (intensity = inferred fluence rate at which filaments change their alignment from parallel to perpendicular to the E vector of plane polarised light).

Nebel (1968) obtained an action spectrum for the phototropic response of chloronemal apical cells of *Physcomitrium turbinatum*. He obtained a peak response with wavelengths around 730 nm, the absorbance peak for phytochrome in its P_{FR} form.

5. CAULONEMAL APICAL CELL PHOTOTROPISM.

In *P. patens*, extension of caulonemal filaments occurs in a similar way to that of chloronemal filaments, by the apical extension of the filament apical cell. Caulonemal apical cells extend 10 – 15 times more rapidly than chloronemal apical cells, and also differ from them, in growing in darkness. Caulonemal filaments are therefore able to function as the adventitious phase of protonemal growth, and allow the rapid colonisation of the soil surface. The polarity of extension of the caulonemal apical cell is sensitive to both light and gravity. In low intensities of red light, apical cell extension occurs towards the light source, but at high intensities, growth is perpendicular (Cove and Lamparter, 1998). No detailed studies on the phototropic response of caulonemal apical cells in this species have been published, but preliminary experiments (Cove et al, 1978) indicated that phytochrome was involved in light perception. In a 23 h darkness, 1 h red light cycle, caulonemal apical cells show a phototropic response. However, when grown in a cycle comprising 22.75 h darkness, 1 h red light, 0.25 h far red light, no phototropic response was observed and

caulonemal apical cells instead showed the characteristic gravitropic response of upward growth, normally shown only in darkness.

It is not possible to differentiate morphologically between chloronemal and caulonemal filaments in the protonemata of *Ceratodon purpureus*. Protonemal apical cells of rapidly-growing axial filaments resemble caulonemal apical cells of *P. patens*, in responding to the direction of both light and gravity, but in contrast to *P. patens*, the phototropic response to red light, is towards the light source in both low and high light intensities. Extensive studies of the phototropic response have been carried out which demonstrate a clear involvement of phytochrome in this phototropic response (Hartmann et al, 1983). Filaments that have been growing in darkness, were exposed to light of different wavelengths and different intensities at right angles to the gravity vector. The angle of growth adopted by the filament was measured and dose response curve constructed. The resulting action spectrum for eliciting the positive phototropic response shows a clear peak at 660 nm. The phototropic response can be reversed if the red light treatment is immediately followed by exposure to far red light. The peak wavelength for far-red reversal is 730 nm. Protonemal apical cells of *C. purpureus* that have been grown in darkness and then illuminated from the side, show a rapid redistribution of the intra-cellular calcium gradient before any morphological change in cell polarity can be observed (Hartmann & Weber, 1988). The highest calcium concentration shifts from the cell apex to towards the light source. Following lateral illumination with red light, early re-organisation of the actin cytoskeleton in protonemal apical cells is also observed (Meske & Hartmann, 1995). This work also showed that anti-microtubule drugs (monensin and colchicines) did not inhibit the phototropic response, and so provided evidence that the microtubule cytoskeleton was not involved in phototropism. It therefore seems likely that calcium and the actin cytoskeleton are both involved downstream of phytochrome in mediating the adjustment of cell polarity to light direction.

Experiments employing the ionophore, monensin, reveal a remarkable aspect of the response of protonemal apical cells to lateral illumination (Hartmann & Weber, 1988). Monensin inhibits polar extension, but can be washed out, allowing growth to resume. If a filament that has been growing in darkness, is treated with monensin and then irradiated laterally with red light, no response can occur, because growth is inhibited. If the irradiation is stopped and the monensin is then washed out, a response to the direction of irradiation occurs. The apical cell can therefore retain a memory of the light direction, even if several hours elapse between the termination of irradiation and the washing out of monensin. The response to red light remains far-red reversible, providing evidence that the phytochrome itself, may be involved in the memory process. However, apical cells that are growing in the absence of monensin do not show the same memory of the light input and need to be irradiated continuously to elicit a phototropic response.

Mutants, impaired in their phototropic response, have been isolated in both *P. patens* (Cove et al, 1978) and *C. purpureus* (Lamparter et al, 1996). Aphototropic mutants of *P. patens* involve at least three complementary genes. Mutations in *ptr*A and *ptr*B lead to the pleiotropic loss of the phototropic response of both caulonemal

apical cells and of gametophores (Courtice, 1979). The status of a further mutant (*ptr*4) that shows a loss of the phototropic response of caulonemal apical cells but not of gametophores, is less clear. This strain complements both *ptr*A and *ptr*B mutants, but when crossed to a phototropically-normal strain, failed to show a simple segregation pattern for the phototropism, indicating that its phenotypes may have a genetically-complex basis (Courtice, 1979).

Genetical analysis of aphototropic mutants of *C. purpureus* are less extensive, but physiological analysis is more advanced. Mutants which have lost the phototropic response of their protonemal apical cells are readily isolated by a selective screening procedure (Lamparter et al, 1996). Mutants fall into two categories. Some show pleiotropic effects, including lower chlorophyll levels, that are consistent with them being impaired in phytochrome-mediated functions, thus providing further evidence that phytochrome is the photoreceptor mediating the phototropic response. Phototropism can be restored in such mutants by feeding with biliverdin, a precursor of the phytochrome chromophore (Lamparter et al, 1996), or phycocyanoblin, which may replace the phytochrome chromophore (Lamparter et al, 1997), suggesting that this class of mutant is deficient in the synthesis of the phytochrome chromophore, at a step before biliverdin. Confirmation that this mutant phenotype is the result of mutation in the gene coding for heme oxygenase, was obtained by demonstrating that the phototropic response of a mutant apical cell could be restored by microinjecting DNA coding for heme oxygenase. (Brücker et al, 2000).

The second class of aphototropic mutant of *C. purpureus*, resembles aphototropic mutants of *P. patens*, in showing no abnormalities in other phytochrome-mediated responses. It is therefore likely that this class of mutant (and the mutants of *P. patens*) is impaired in its phototropic response downstream of light perception by phytochrome.

There are no reports of crosses between aphototropic mutants of *C. purpureus* and the result of only one somatic hybridisation experiment has been described (Lamparter et al, 1998a). A mixture of protoplasts from *ptr*116, an aphototropic mutant repairable by biliverdin, and *ptr*103, a non-repairable mutant, were treated to induce protoplast fusion. Among the regenerants recovered, some showed a wild-type phototropic response. Since no selective markers were available, the genome size of a wild-type regenerant was estimated and found to be double that of either of the *ptr* strains, confirming that it was a somatic hybrid, and therefore that complementation occurred between two classes of aphototropic mutants.

6. CAULONEMAL APICAL CELL GRAVITROPISM

The gravitropic response of protonemal filaments has been studied in a number of moss species including *Funaria hygrometrica* (Schwuchow et al, 1995), *Pottia intermedia* (Chaban et al, 1998), and, in most detail, *C. purpureus* and *P. patens*. The apical cells of caulonemal filaments of wild-type strains *P. patens* (Cove et al, 1978) and of protonemal filaments of wild-type strains of *C. purpureus* (Hartmann et al, 1983, Walker and Sack, 1990) grow upwards in darkness (classically described as

showing negative gravitropism) (see figure 2).

Strong growth of filaments in darkness is only obtained on media supplemented with either glucose or sucrose. Apical extension of caulonemal filaments of *P. patens* occurs at about 30 - 40 µm/h at 25° (Knight and Cove, 1991). Growth of *C. purpureus* filaments in darkness is more vigorous, and apical extension occurs at about the same rate for wild type 3, but somewhat faster for wild type 4 (see figure 2). The gravitropic response is inhibited by exposure to light, even at low

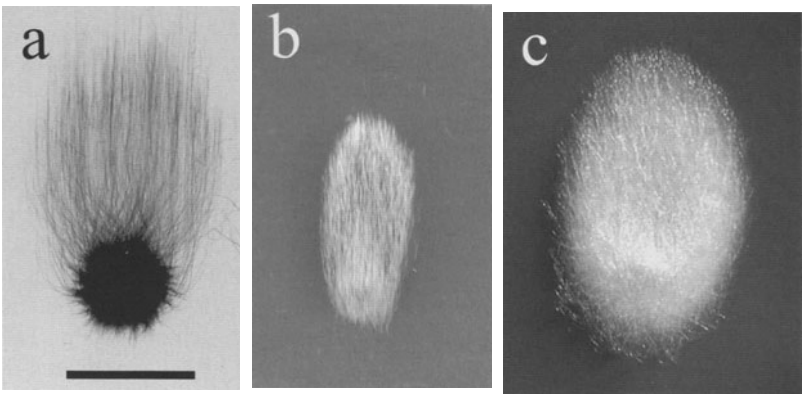

Figure 2. Negatively-gravitropic growth of protonemata in darkness.

a. *P. patens* Gransden Wild Type (Ashton & Cove, 1977) grown on minimal medium supplemented with sucrose. After initial culture for 14d in white light, the Petri dish was placed vertically and transferred to darkness for a further 21d
b. *C. purpureus* Wild Type 3 (Hartman et al, 1983) cultured as in a, but transferred to darkness for 14d after incubation in light for 3d.
c. *C. purpureus* Wild Type 4 (Hartman et al, 1983) cultured as in b.

Scale bar for = 10 mm

intensities (see below). When dark-grown caulonemal filaments of *P. patens* are re-oriented by 90° with respect to the gravity vector, a 90° response gravitropic is achieved in about 72 h (see figure 3). Protonemal filaments of *C. purpureus* respond more rapidly, a full response to 90° reorientation being achieved in about 24 h (Walker and Sack, 1990). Time lapse video microscopy of filaments responding to reorientation to the gravity vector, is possible by filming in long wavelength light (>850 nm) and using an infra-red sensitive camera (Knight and Cove, 1991, Young and Sack, 1992, Schwuchow et al, 1995).

These studies reveal that the negative gravitropic response is not continuous. Immediately following reorientation, growth is towards the gravity vector rather than away and positively gravitropic growth also occurs during each mitotic division of the apical cell. These are not the only circumstances under which a positively gravitropic

response of a wild-type apical cell is observed. A positive response occurs in *P. patens* filaments, immediately upon transfer to a clinostat (Knight and Cove, 1991) and

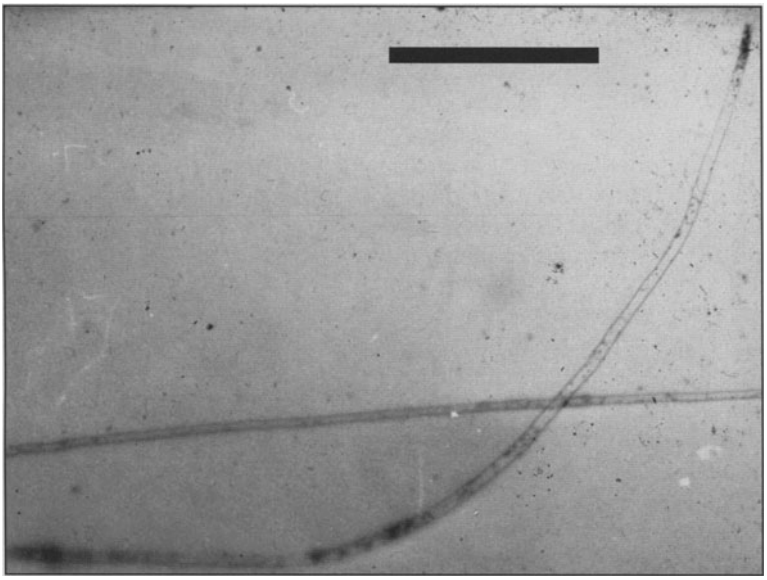

Figure 3. Growth response of a caulonemal filament of *P. patens* to 90° reorientation with respect to the gravity vector.. Scale bar = 400 μm

in *C. purpureus* filaments, when grown in monochromatic blue light (Lamparter *et al*, 1998b). The significance of these positively-gravitropic response is not understood, but it has been suggested that the partial disassembly of the microtubular cytoskeleton during mitoses, may interfere with the response to gravity at those times (Knight and Cove 1991). Evidence for the role of the microtubular cytoskeleton in gravitropism is provided by inhibitor studies (Schwuchow *et al*, 1990; Meske and Hartmann, 1995). Whereas the anti-microtubule drug, oryzalin does not inhibit the phototropic response of *C. purpureus* protonemal apical cells, this drug inhibits the gravitropic response.

There is a marked zonation of plastids along the long axis of apical cells of protonemal filaments of *C. purpureus* (Walker & Sack, 1990) and of caulonemal filaments *P. patens* (Knight & Cove, 1991) and *F. hygrometrica* (Schwuchow et al, 1995). In *C. purpureus*, there is a zone of plastids at the immediate tip, which is absent in the other two species. In all three species, moving away from the cell apex, there is a plastid-free zone. Next there is a zone of amyloplasts which exhibit clear gravity-stimulated sedimentation in *C. purpureus* (Walker & Sack, 1990) and *F. hygrometrica* (Schwuchow et al, 1995), and finally a zone of containing non-sedimenting plastids and a vacuole. Plastid sedimentation in *P. patens* is not obvious Knight & Cove, 1991) but some movement is detectable (Schwuchow *et al*, 1995). It is attractive to propose that the sedimenting zone of amyloplasts is the site of gravity perception by apical

cells, but studies of regenerating protoplasts reveal that emerging filaments become graviperceptive before plastid zonation or sedimentation can be observed (Wagner & Sack, 1998).

In a preliminary study of the interaction between phototropism and gravitropism in *P. patens* (Cove *et al*, 1978), protonemal growth in light was found to be "completely indifferent to gravity". Jenkins *et al* (1986) found that in intensities of monochromatic red light above 200 nmol quanta m^{-2} s^{-1} no sensitivity to gravity could be detected. When irradiated at right angles to the gravity vector, with monochromatic red light at an intensity of 90 nmol quanta m^{-2} s^{-1}, caulonemal apical cells appear to make a choice between positive phototropism and negative gravitropism, resulting in filaments that switch from time to time, between upwards growth and growth towards the light, but irradiation at an intensity of 60 nmol quanta m^{-2} s^{-1}, results in growth at 45° to both the light direction and gravity (Cove & Knight, 1987, Jenkins & Cove, unpublished data). These studies suggest that there is competition between gravitropism and phototropism, and that phototropism is the stronger morphogenetic input. However, studies of aphototropic mutants of both *P. patens* and *C. purpureus* show that the interaction of gravitropism with light, is not simply competitive. Thus aphototropic mutants of *P. patens*, which have a wild-type gravitropic response in darkness, do not respond to gravity in light levels above 1 μmol quanta m^{-2} s^{-1} (Cove & Knight, 1987). Mutants of the second class of aphototropic mutant of *C. purpureus*, (impaired downstream of phytochrome - see above), show a similar lack of gravitropism in light. On the other hand, *C. purpureus* mutants that are impaired in the synthesis of the phytochrome chromophore, continue to respond to gravity even in high light levels (Lamparter *et al*, 1996), providing strong evidence that gravitropism is actively switched off in the light, by way of phytochrome.

Mutants showing an abnormal response to gravity (*gtr* mutants) , have been identified in both *P. patens* and *C. purpureus*. Genetic analysis of *P. patens gtr* mutants has been carried out using somatic hybridisation (Jenkins *et al*, 1986). This initial study revealed that at least three genes are involved. *gtr*A mutants are impaired in their response to gravity but even those most extremely affected, still show a tendency to grow upwards. These mutants show a pleitropic effect on gametophore development, having more rounded leaf cells and consequently more fleshy leaves. Mutants isolated initially as a result of their having this leaf morphology, have been shown subsequently to be impaired in their gravitropic response, confirming this pleiotropy. A second gene was identified by a single mutant, *gtr*B1. The morphology of the caulonemal apical cells of this mutant was abnormal with no plastid zonation observable (Jenkins *et al*, 1986). Unfortunately, this mutant was later lost and no similar mutants have since been reported. A third gene was identifiedin in the original study, again by a single mutant, *gtr*C5. This mutant produced caulonemal apical cells that responded positively to gravity, *i.e.* grew downwards. A subsequent more extensive genetic analysis (Knight, 1987) included more mutants. This identified a further gene, also represented by a single mutant, *gtr*D8, which like *gtr*A mutants showed a pleiotropic effect on leaf morphology. This study included more *gtr*C alleles, all with caulonemal apical cells showing a positive gravitropic response. Genetic analysis showed that the *gtr*C7

mutant segregated in a regular manner, confirming that this phenotype arises as a result of mutation in a single gene.

All mutants of *P. patens* showing an abnormal gravitropic response of the caulonemal apical cells, produced gametophores showing the wild-type negatively gravitropic response, contrasting with mutants affected in phototropism, where the response of both caulonemal apical cells and of gametophores was affected (see above). No mutants having gametophores impaired in their response to gravity have been reported.

Mutants of *C. purpureus* with an altered protonemal response to gravity are readily selected as a by-product of the isolation of phototropically-abnormal strains (Lamparter *et al*, 1996; Wagner *et al*, 1997). Mutants resembling *P. patens gtr*A in showing a weak or very weak negatively gravitropic response appear to be less common (Cove, unpublished data), but mutants responding positively to gravity are more easily identified. The first of these to be examined physiologically was designated *wwr*1 (for wrong way response), although subsequent genetic analysis suggests that a similar mutant may be an allele of a gene that can mutate to give a near-agravitropic phenotype (see below). The *wwr*1 mutant was found to have a gravitropic response which mirrored that of the wild type, in that initially and during nuclear division it grew upwards, and at other times down (Wagner *et al*, 1997). The protonemal apical cells of the *wwr*1 mutant show similar but not identical plastid zonation to the wild type, and plastid sedimentation is obvious in the mutant.

The absence of mutants in both *P. patens* and *C. purpureus* that are completely agravitropic, suggests that there may be some overlap of gene function such that knock out of any one gene does not abolish sensitivity to the gravity vector completely. The recent report that *C. purpureus*, grown in microgravity, has a characteristic symmetrical spiral growth (F. Sack unpublished data, see - http://spaceresearch.nasa.gov/general_info/16jul_firemoss_lite.html), suggest that this is the predicted phenotype of a mutant strain completely insensitive to gravity and a further mutant hunt beginning with a near-agravitropic strain, might identify a further gene involved in gravisensing.

Genetic analysis of gravitropic mutants of *C. purpureus*, is in progress (Cove, unpublished data). Somatic hybrids are selected using transgenic antibiotic resistant strains obtained following transformation. Hybrids may be selected between strains having resistances to different antibiotics, by counter-selecting for both antibiotics. Figure 4 shows a series of hybrids obtained in this manner. The control (homozygous) hybrids, show the phenotypes of the wild type and gravitropic mutants involved, one of which is almost completely impaired in its response (*gtr*406), and the other (*gtr*445) showing a reversed positively-gravitropic response.

The hybrids between *gtr*406 and wild type, and between *gtr*445 and wild type, both show the wild-type response, confirming that both mutant phenotypes are recessive. The *grtr*406/*gtr*445 does not however show a wild-type response as would have been predicted had the two phenotypes been the result of mutation in different genes. The phenotype of the *gtr*405/*gtr*446 hybrid may indicate that a single gene can mutate to give rise to alleles which result in either loss or reversal of gravitropism, in

which case the positively-gravitropic phenotype presumably involves only a partial loss of gene function. Alternatively, *gtr*405 may be a double mutant, agravitropism being epistatic to the reversed response. Analysis involving more mutant strains, which is currently in progress, may allow these possibilities to be distinguished.

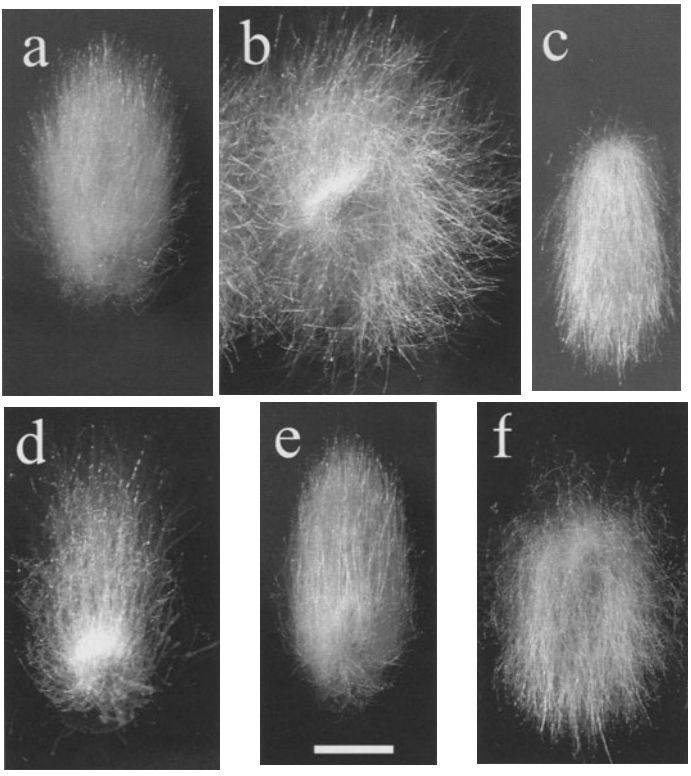

Figure 4. Growth of somatic hybrids of *C. purpureus* in darkness.

a	wild type plus wild type	d	*gtr*406 plus wild type
b	*gtr*406 plus *gtr*406	e	*gtr*445 plus wild type
c	*gtr*445 plus *gtr*445	f	*gtr*406 plus *gtr*445

Somatic hybrids were obtained by fusion of protoplasts from two transgenic strains containing genes conferring resistance to different antibiotics, following by selection for strains resistant to both antibiotics (Cove, unpublished data). Scale bar: 10 mm

7. GAMETOPHORE GRAVITROPISM AND PHOTOTROPISM

There are no reports of extensive studies on gametophore tropisms. In *P. patens*, the development of new gametophores does not occur in darkness, but gametophores present upon transfer to darkness, etiolate and grow upwards (Cove *et al*, 1978). As with the response of caulonemal apical cells, even low levels of light inhibit the gravitropic response of gametophores, with those of the wild type growing toward the light and those of aphototropic mutants oriented randomly. It therefore seems likely that gametophore gravitropism is also actively switched of by light.

8. FUTURE PROSPECTS

Studies to date have shown that moss development provides outstanding material for the study of polarity at both the level of the individual cell and in multicellular structures. To advance these studies, more detailed investigation of the cellular events involved in the establishment and modification of polarity, using modern cell biological techniques, need to be combined with molecular genetic studies of the genes involved. If progress is forthcoming, moss development has the potential to make landmark contributions to our understanding of biological polarity.

Public access to an extensive EST database, as well as libraries of genomic DNA, bacterial artificial chromosomes (BACs), and cDNAs, are now available for *P. patens* (see - www.moss.leeds.ac.uk). These resources provide the possibility of identifying genes of potential interest in relation to polarity, by their homology to genes identified in other systems, including higher plants (see for example Fu *et al*, 2001), the products of which play a role in the establishment and/or maintenance of a polar axis. Coupling these resources with the ability to remove or replace specific genes and/or promoters by homologous recombination (Schaefer, 2001) provides a unique opportunity in plants to localize and to determine the function of specific gene products in cell polarity. The ability to target transgenes to specific genomic loci (Schaefer, 2001), the disruption of which does not have any noticeable effect on the development of *P. patens*, allows the opportunity to assess in stable lines of *P patens,* the results of ectopic expression of genes of interest. Effects on cell polarity of the over-expression of genes, and the expression of anti-sense and/or RNAi constructs can be assessed, independent of the complexities of interpreting the effects of transgenes in different chromatin environments. Similarly, the replacement of wild-type genes with alleles of specific interest will allow an assessment of the effects of the allelic substitution itself which will not be subject to effects due to the chromosomal environment. This is not yet possible in *Arabidopsis thaliana*, where for example, Fu *et al.* (2001) have shown specific effects of At ROP1 alleles on pollen tube growth, but due to the nature of transformation in *A. thaliana*, each of these alleles is located in a different chromosomal environment, not that of its wild type counterpart. A very similar homolog to the *A. thaliana* ROP1 gene is present in the *P. patens* genome. The

tools now exist in *P. patens* to study the effects of these same allelic substitutions on filament tip growth at the same (wild-type) chromosomal location.

9. REFERENCES

Ashton, N.W. & Cove, D.J. 1977. The isolation and preliminary characterisation of auxotrophic mutants of the moss, *Physcomitrella patens*. Molecular and General Genetics 154, 87-95.

Brücker, G., Zeidler, M., Kohchi, T., Hartmann, E. & Lamparter, T. 2000. Phytochrome-chromophore deficient mutants of the moss *Ceratodon purpureus* can be rescued by microinjecting heme oxygenase genes. Planta 210, 529-535

Courtice, G.R.M. 1979. Developmental genetic studies of *Physcomitrella patens*. Ph. D. thesis, University of Cambridge, England

Chaban, Ch.I., Kern, V.D., Ripetsky, R.T., Demkiv, O.T. & Sack, F.D. 1998. Gravitropism in caulonemata of the moss *Pottia intermedia*. Journal of Bryology 20, 287-299

Cove, D.J., Schild, A., Ashton, N.W. & Hartmann, E. 1978. Genetic and physiological studies of the effect of light on the development of the moss, *Physcomitrella patens*. Photochemistry and Photobiology 27, 249-254.

Cove, D.J. & Knight C.D. 1987. Gravitropism and phototropism in the moss, *Physcomitrella patens*. In "Developmental Mutants of Higher Plants" pp 181-196. Editors: H. Thomas and D. Grierson. Published: Academic Press, London, England 1987.

Cove, D.J., Quatrano, R.S. & Hartmann, E. 1996. The alignment of the axis of asymmetry in regenerating protoplasts of the moss, *Ceratodon purpureus*, is determined independently of axis polarity. Development 122, 371 - 397.

Cove, D.J. & Lamparter, T. 1998. The role of light in the regulation of moss development. In "Microbial Responses to Light and Time" (Eds.: M.X.Caddick, S. Baumberg, D.A.Hodgson and M.K.Phillips-Jones). pp. 125 – 141. Cambridge University Press, London

Cove, D.J., Hope, I.A. & Quatrano, R.S. 1999. Polarity in biological systems. *In* "Development: Genetics, Epigenetics and Environmental Regulation". (Eds.: V.E.O.Russo, D.J.Cove, L.Edgar, R.Jaenisch and F.Salamini). pp. 507 - 524. Springer-Verlag, Heidelberg.

Fu, Y., Wu, G. & Yang, Z. .2001. Rop GTPase-dependent dynamics of tip-localized F-actin controls tip growth in pollen tubes. Journal of Cell Biology 152, 1019-1032.

Hartmann, E., Klingenberg, B. & Bauer, L. 1983. Phytochrome-mediated phototropism in protonemata of the moss, *Ceratodon purpureus* Brid. Photochemistry and Photobiology 38, 599–603.

Hartmann, E. & Weber, M. 1988. Storage of the phytochrome-mediated phototropic stimulus of moss protonemal tip cells. Planta 175, 39–49.

Jaffe, L. & Etzold, H. 1965. Tropic responses of *Funaria* spores to red light. Biophysical Journal 5, 715–742.

Jenkins, G.I., Courtice, G.R.M. & Cove, D.J. 1986. Gravitropic responses of wild-type and mutant strains of the moss, *Physcomitrella patens*. Plant, Cell and Environment 9, 637-644.

Jenkins, G.I. & Cove, D.J. 1983a. Phototropism and polarotropism of primary chloronemata of the moss *Physcomitrella patens*: responses of the wild-type. Planta 158, 357-364.

Jenkins, G.I. & Cove, D.J. 1983b. Phototropism and polarotropism of primary chloronemata of the moss *Physcomitrella patens*: responses of mutant strains. Planta 159, 432-438.

Knight, C.D. 1987. Gravitropism in the moss, *Physcomitrella patens*. Ph.D. thesis, University of Leeds, England.

Knight, C.D. & Cove, D.J. 1991. The polarity of gravitropism in the moss *Physcomitrella patens* is reversed during mitosis and after growth on a clinostat. Plant, Cell and Environment 14, 995-1001.

Lamparter, T., Esch, H., Cove, D., Hughes, J. & Hartmann, E. 1996. Aphototropic mutants of the moss *Ceratodon purpureus* with spectrally normal and with spectrally dysfunctional phytochrome. Plant, Cell and Environment 19, 560-568.

Lamparter, T., Esch, H., Cove, D.J. & Hartmann, E. 1997. Phytochrome control of phototropism and chlorophyll accumulation in the apical cells of protonemal filaments of wildtype and an aphototropic mutant of the moss *Ceratodon purpureus*. Plant and Cell Physiology 38, 51-58.

Lamparter, T., Brücker, D., Esch, H., Hughes, J., Meister, A. & Hartmann, E. 1998a. Somatic hybridisation with aphototropic mutants of the moss *Ceratodon purpureus*: genome size, phytochrome photoreversibility, tip-cell phototropism and chlorophyll regulation. Journal of Plant Physiology 153, 394-400.

Lamparter, T., Hughes, J., Hartmann, E. 1998b. Blue light- and genetically-reversed gravitropic response in protonemata of the moss *Ceratodon purpureus*. Planta 206, 95-102.

Meske, V. & Hartmann, E. 1995. Reorganisation of microfilaments in protonemal tip cells of the moss, *Ceratodon purpureus*. Protoplasma 188, 58–68.

Schaefer, D.G. 2001. Gene targeting in *Physcomitrella patens*. Current Opinion in Plant Biology 4, 143-150.

Schildt, A. 1981. Untersuchungen zur sporenkeimung und protonemaentwicklung bei dem laubmoss *Physcomitrella patens*. Ph. D. thesis, University of Mainz, Germany.

Schwuchow, J.M., Kim, D. & Sack, F.D. 1995. Caulonemal gravitropism and amyloplast sedimentation in the moss *Funaria hygrometrica*. Canadian Journal of Botany 73, 1029-1035.

Wagner, T.A., Cove, D.J. & Sack, F.D. 1997. A positively gravitropic mutant mirrors the wild-type protonemal response in the moss *Ceratodon*. Planta 202, 149 - 154.

Wagner, T.A. & Sack, F.D. 1998. Gravitropism and gravimorphism during regeneration from protoplasts of the moss *Ceratodon purpureus* (Hedw.) Brid. Planta 205, 352-358.

Walker, L.M. & Sack, F.D. 1990. Amyloplasts as possible statoliths in gravitropic protonemata of the moss *Ceratodon purpureus*. Planta 181, 71-77.

Young, J.C. & Sack, F.D. 1992. Time-lapse analysis of gravitropism in *Ceratodon* protonemata. American Journal of Botany 79, 1348-1358.